Draft Systematic Review

Number xx

Strategies to Improve Mental Health Care for Children and Adolescents

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No.

Prepared by:

Investigators:

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by an Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. xxx-xxxx-xxxxx). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies may not be stated or implied.

This report may periodically be assessed for the currency of conclusions. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program Web site at www.effectivehealthcare.ahrq.gov. Search on the title of the report.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact [insert program email address].

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: To Be Added to Final Version.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new health care technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

We welcome comments on this systematic review. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Richard G. Kronick, Ph.D.
Director
Agency for Healthcare Research and Quality

Arlene Bierman, M.D., M.S.
Director, Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director, EPC Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Christine Chang, M.D., M.P.H.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: To Be Added to Final Version.

Key Informants

In designing the study questions, the EPC consulted a panel of Key Informants who represent subject experts and end-users of research. Key Informant input can inform key issues related to the topic of the technical brief. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who participated in developing this report follows:

To Be Added to Final Version.

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who participated in developing this report follows:

To Be Added to Final Version.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers follows:

To Be Added to Final Version.

Strategies to Improve Mental Health Care for Children and Adolescents

Structured Abstract

Objectives. To increase knowledge about the effectiveness of quality improvement (QI), implementation, and dissemination strategies that seek to improve the mental health care of children and adolescents (Key Question [KQ] 1); to examine harms associated with these strategies (KQ 2); and to determine whether effectiveness or harms vary in subgroups based on system, organizational, practitioner, or patient characteristics (KQ 3).

Data sources. MEDLINE®, Cochrane Library, PsycINFO, and CINAHL (Cumulative Index to Nursing and Allied Health Literature), gray literature, additional studies from reference lists and technical experts.

Review methods. Two trained reviewers selected, extracted data from, and rated the risk of bias of relevant trials and cohort studies. We synthesized the data qualitatively and graded strength of evidence for each outcome. For all outcomes, we present relative risks or mean differences, with confidence intervals, whenever calculable. For outcomes with multiple measures, we present forest plots.

Results. We found 15 studies from 14 publications on KQ 1 (overall effectiveness), one on KQ 2 (harms), and four on KQ 3 (moderators of effectiveness), of which 11 were randomized controlled trials, 2 controlled clinical trials, 1 quasi-experiment, and 1 ecological aggregate. The strategies included in this review were complex, heterogeneous, and difficult to categorize. Six of our studies (reported in 5 publications) tested strategies that spanned multiple categories of our original classification scheme that included implementation, dissemination, or quality improvement. This overlap prompted us to use a different system, based on the Cochrane Review Group's Effective Practice and Organisation of Care taxonomy, to categorize specific components of each strategy. The categorization of components enabled us to classify strategies as professional training (i.e., strategies that comprised only professional components, n=6 studies in 5 publications) or financial or organizational change (i.e., strategies that comprised at least one financial or organizational component, n=9 studies), although all strategies included more than one component so synthesis of findings was difficult. We also identified multiple active components (in 11 studies) or the single active component (in 4 studies) in each strategy.

We graded 14 outcomes for professional training and 16 for financial or organizational change. We graded the strength of over half of these outcomes as insufficient or low for no benefit. We found evidence that a majority of strategies had at least some evidence of effectiveness. Ten studies reported in 9 publications (i.e., 9 strategies) had at least one outcome rated as low for benefit, and one study had a single outcome rated as moderate for benefit. We were unable to judge the overall potential for harms associated with these strategies that may mitigate benefits based on the single included study with information on harms (KQ 2). The available evidence from four studies on two moderators does not permit us to make general conclusions about the conditions under which these strategies might work optimally (KQ 3).

Conclusions. The evidence does not permit us to have a high degree of confidence about the efficacy of any one strategy because we generally found a single study testing each strategy. We found the strongest or most consistent evidence of benefit for strategies with that provided financial benefits to practitioners to maintain fidelity to an evidence-based practice (EBP) and strategies that provide professional training to improve access to EBPs.

Contents

Tables

Figures

Appendixes

Executive Summary

Background

Approximately one in five children and adolescents living in the United States has one or more mental, emotional, or behavioral health disorders according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) criteria in any given year.[1] These disorders contribute to problems with family, peers, and academic functioning. They may exacerbate coexisting conditions (including other mental and substance use disorders and chronic health conditions) and may reduce quality of life. They also increase the risk of involvement with the criminal justice system and other risk-taking behaviors and suicide.[2]

Several key publications in the mid- to late-1990s suggested that usual care in children's mental health had, at best, no[3] and sometimes harmful effects.[4] Since then, the evidence base for pediatric mental health interventions that target mood disorders, anxiety disorders, disruptive behavior disorders, psychotic disorders, eating disorders, and substance use disorders has grown.[5, 6]

Despite advances in the evidence base,[5, 7] outcomes for children with mental health problems remain suboptimal because of issues with access to care and the failure of systems and providers to adopt established quality improvement (QI) strategies and interventions with proven effectiveness (e.g., evidence-based practices [EBPs]). Studies using nationally representative data on U.S. adolescents show that only approximately one in five children with mental health problems receives services, and only one-third of treatment episodes are considered minimally adequate (at least four visits with psychotropic medication or at least eight visits without psychotropic medication).[8-10] The current health care system continues to provide fragmented care to children and adolescents in numerous uncoordinated systems, rendering inefficient the delivery of needed services.[11] Moreover, clinicians (particularly primary care practitioners) may lack the time, knowledge, or training to identify and treat or refer mental health problems appropriately.[12]

Given the gap between observed and achievable processes and outcomes, the next step is the adoption of QI strategies and the development of strategies to implement or disseminate these interventions. Such strategies target changes in the organization and delivery of mental health services.[13, 14] They seek to improve the quality of care and patient outcomes by closing the gap between research evidence and practice.[15-17]

The Cochrane Collaboration's Effective Practice and Organisation of Care (EPOC) Group studies complex strategies designed to improve health care professionals' practice and the organization of health care services. The EPOC group's taxonomy classifies these strategies as having one or more professional, financial, organizational, and regulatory components. These strategies typically include various forms of continuing education for providers; quality assurance projects; and financial, organizational, or regulatory interventions that can affect the ability of health care professionals to deliver services more effectively and efficiently.

The ultimate goal of the strategies included in this review is to improve patient health and service utilization outcomes for children and adolescents with mental health problems. Intermediate outcomes in this context include changes to health care systems, organizations, and practitioners that provide mental health care.

Scope and Key Questions

As reflected in our Key Questions (KQs) and analytic framework below, we have three primary aims for this review. First, we will increase knowledge about the effectiveness of dissemination, implementation, or QI strategies that seek to improve the mental health care of children and adolescents. Second, we will examine harms associated with these strategies. Third, we will attempt to determine whether effectiveness or harms vary in subgroups based on system, organizational, practitioner, or patient characteristics.

Key Questions

KQ 1: What is the effectiveness of quality improvement, implementation, and dissemination strategies employed in outpatient settings by health care practitioners, organizations, or systems that care for children and adolescents with mental health problems to improve:

a. intermediate patient, provider, or system outcomes
b. patient health and service utilization outcomes?[1]

KQ 2: What are the harms of these mental health strategies?

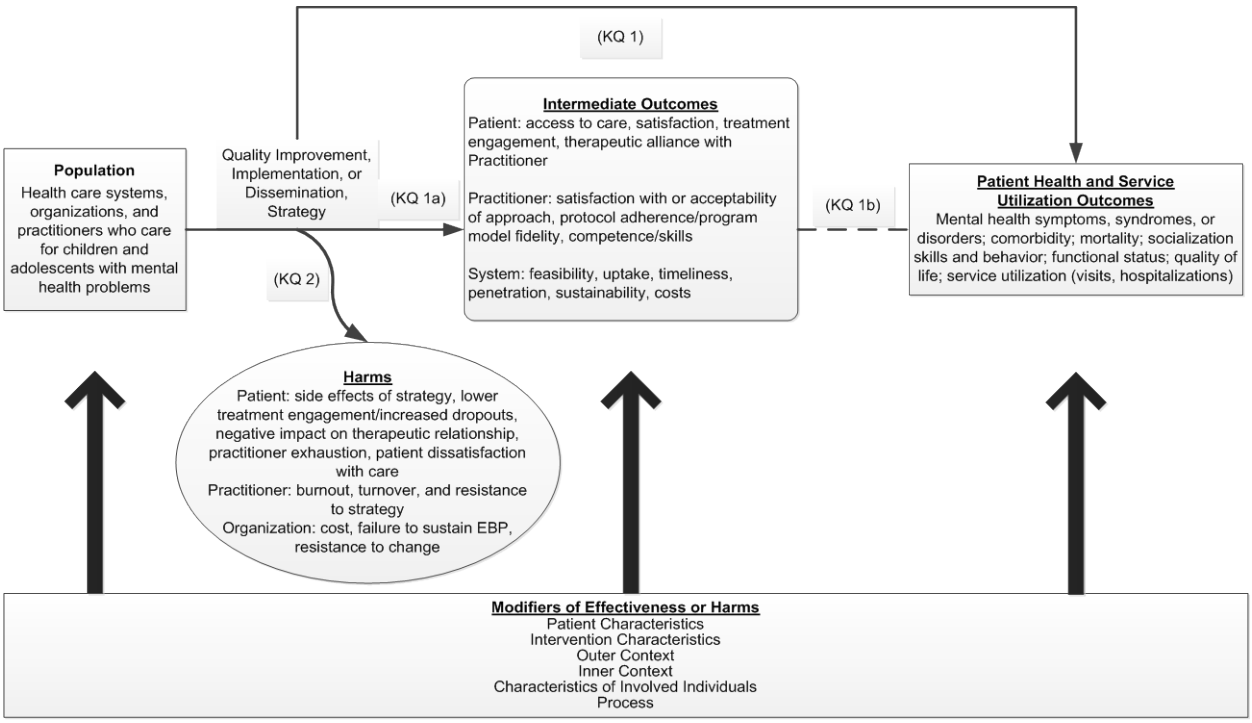
KQ 3: Do characteristics of the child or adolescent or contextual factors (e.g., characteristics of patients, practitioners, organizations, or systems; intervention characteristics; setting; or process) modify the effectiveness or harms of strategies to improve mental health care and, if so, how?

Analytic Framework

Figure A depicts the patient populations, interventions, comparators, outcomes, and timing of outcomes assessment (PICOTs) and KQs in relation to these PICOTs.

[1] We revised KQ 1 and the outcome specified in our protocol slightly for clarity. We replaced the term "health care providers" with "health care practitioners" to indicate that this particular phrase refers to individuals rather than systems or institutions. We also replaced "final outcomes" with "patient health and service utilization outcomes" for clarity.

Figure A. Analytic framework for strategies to improve mental health care in children and adolescents



Populations, Interventions, Comparators, Outcomes, Timing, and Setting

We specified our inclusion and exclusion criteria based on the PICOTS identified through the topic refinement exercise. We included QI, implementation, and dissemination strategies that targeted systems, organizations, or practitioners of mental health care to children and adolescents ages 18 years of age or younger, who were already experiencing mental health symptoms. We did not include strategies such as the implementation of educational interventions for reading disorders. We also limited our review of implementation strategies to those focusing on EBP interventions. For defining EBPs, we relied on the minimum requirements set forth by the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-based Programs and Practices (NREPP) (www.nrepp.samhsa.gov). These criteria specify that the intervention needs to have produced one or more positive behavioral outcomes in at least one study using an experimental or quasi-experimental design with results published in a peer-reviewed journal or similar publication. In addition, implementation materials, training and support resources, and quality assurance procedures for these interventions need to be ready for use by the public.

We use the term "strategy" to reference the total sum of components used to target health care systems and/or practitioners to improve the quality of care for children and adolescents with mental health problems. We use the term "intervention" to denote a specific EBP used as part of a strategy.

Because strategies tended to be complex in nature and the number and types of components that varied between the treatment arm and comparison group arm differed by study, we also recorded components of each strategy by study arm according to the EPOC taxonomy.[18] Because many of the comparison groups also contained several components, we marked the components

contained in each study arm of each study. This allowed us to fully describe the numerous components that were being combined and tested in each strategy, as well as enabled us to determine whether the study arms differed by a single or multiple components.

We required each included study to report at least one intermediate outcome in a minimum of one of three major categories: (1) practitioner intermediate outcomes (satisfaction, adherence, fidelity, competence), (2) system intermediate outcomes (feasibility, uptake, timeliness, penetration, sustainability, costs), and (3) patient intermediate outcomes (access to care, satisfaction, engagement, therapeutic alliance). This helped to ensure that each included study demonstrated impact based on its stated goals of improving quality or implementing or disseminating evidence-based interventions. We also required each study to report at least one patient health or service utilization outcome (change in mental health status, comorbid conditions, mortality, socialization skills and behavior, functional status, quality of life, service utilization) if the strategy was not implementing or disseminating an EBP (i.e., an intervention with proven effectiveness).

For all KQs, we excluded study designs without comparison groups to ensure that our pool of included studies provided evidence on the causal link between the strategy and outcomes. We also required that the comparator enabled examination of the strategy effectiveness. That is, we excluded studies in which the strategy (system, organizational, practitioner targets) and the intervention being tested both differed between groups, because the effectiveness of the QI, implementation, or dissemination strategy could not be isolated from the baseline intervention effects.

Our exclusion of non–English-language studies is based on limitations of time and resources. However, we examined English language abstracts of non–English-language studies to assess the potential size of the literature that would be missed through this approach.

Table A. Populations, interventions, comparators, outcomes, timing, and settings

Category	Inclusion and Exclusion Criteria and Relevant Factors for Study Abstraction
Population	*Inclusion criteria:* Health care systems, organizations, and practitioners that care for children and adolescents or mixed (child and adult) populations with mental health problems Health care systems, organizations, and practitioners that care only for adults 18 years of age or older *Exclusion criteria:* Health care systems, organizations, and practitioners that care for children and adolescents with only developmental disorders *Relevant factors:* Patient characteristics, such as age, gender, race and ethnicity, cognitive ability, diagnosis and severity of mental health problem, coexisting conditions, and cotreatments

Table A. Populations, interventions, comparators, outcomes, timing, and settings (continued)

Category	Inclusion and Exclusion Criteria and Relevant Factors for Study Abstraction
Interventions (Strategies)	*Inclusion criteria:* • Quality improvement strategies (e.g., strategies targeting systems and practitioners of mental health care to children and adolescents with the goal of improved quality of care) • Implementation strategies (e.g., strategies to integrate EBP interventions that meet NREPP inclusion criteria with the goal of changing practice patterns) • Dissemination strategies (e.g., strategies to enhance the adoption and the implementation of evidence-based interventions that meet NREPP inclusion criteria) *Exclusion criteria:* Interventions targeting only patients, only drug interventions, and interventions not otherwise described in inclusion criteria *Relevant factors:* Intervention characteristics, such as complexity; manualized or not; intensity, frequency or duration; and adjustment of intervention to fit context Process characteristics, such as fidelity to the planned strategy, fidelity to the EBP, use of champions, and supervision or oversight Characteristics of involved individuals such as type, knowledge, beliefs, self-efficacy, leadership, education, certifications, and years of practice of practitioners or characteristics of parents/caregivers Other components, such as length of followup
Comparator	*Inclusion criteria:* Any control strategy, including usual care or different variants of the same intervention None
Outcomes	*Inclusion criteria:* **Intermediate outcomes (at least one intermediate outcome is required for KQs 1, 3)** • Patient – access to care – satisfaction – treatment engagement – therapeutic alliance with practitioner • Practitioner – satisfaction with or acceptability of approach – protocol adherence/program model fidelity – competence or skills • System or organization – feasibility – uptake – timeliness – penetration – sustainability – resources (including costs) **Patient health and service utilization outcomes (at least one of these outcomes is required for KQs 1 and 3 unless the strategy uses an intervention that is an EBP)** • Change in mental health status, including symptom change, response, remission, relapse, and recurrence • Coexisting physical health conditions, substance use problems, developmental disorders, other mental health problems • Mortality • Socialization skills and behavior • Functional status • Quality of life

Table A. Populations, interventions, comparators, outcomes, timing, and settings (continued)

Category	Inclusion and Exclusion Criteria and Relevant Factors for Study Abstraction
	Service utilization (e.g., visits, hospitalizations) Harms of strategy • Patient – lower treatment engagement or more dropouts – negative impact on therapeutic relationship – side effects of evidence-based practice incorporated into strategy (e.g., adverse events, suicidality) – patient dissatisfaction with care • Practitioner – burnout or exhaustion – turnover – resistance to the intervention • System or organization – cost – failure to sustain the evidence-based practice – resistance to change *Exclusion criteria:* *All outcomes not otherwise specified*
Timing of outcome measurement	*Inclusion criteria:* All *Exclusion criteria:* None
Settings	*Inclusion criteria:* Outpatient settings serving children and adolescents with mental health problems (primary care, specialty care, emergency rooms, community mental health centers, integrated care settings, federally qualified health centers, schools, homes) *Exclusion criteria:* Inpatient or residential treatment settings, drug treatment programs, jails, or prisons *Relevant factors:* Outer setting, such as external policy, incentives, availability of alternative care systems; Inner setting or organizational factors, such as type of outpatient setting, structure or size, culture, implementation climate, and readiness of organization for implementation
Geographic setting	*Inclusion criteria:* Countries with a very high human development index (HDI)[19] *Exclusion criteria:* Countries with high, medium, low, or very low HDI
Publication language	*Inclusion criteria:* English *Exclusion criteria:* All other languages

Table A. Populations, interventions, comparators, outcomes, timing, and settings (continued)

Category	Inclusion and Exclusion Criteria and Relevant Factors for Study Abstraction
Study design	*Inclusion criteria:* KQs 1, 3 (benefits) • RCTs • CCTs • Systematic review and meta-analyses • Cohort studies • Interrupted time series • Case-control studies KQs 2, 3 (harms): • RCTs • CCTs • Systematic review and meta-analyses • Cohort studies • Interrupted time series • Case-control studies *Exclusion criteria:* • Case series • Case reports • Nonsystematic reviews • Cross-sectional studies • Before and after studies without time series data • Other designs without a control or comparison group
Publication type	*Inclusion criteria:* Any publication reporting primary data *Exclusion criteria:* Publications not reporting primary data

CCT = controlled clinical trial; EBP = evidence-based practice; HDI = Human Development Index; KQ = Key Question; NREPP = National Registry of Evidence-based Programs and Practices; RCT = randomized controlled trial.

Methods

The methods for this systematic review follow the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* from AHRQ (available at http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm). The review uses the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist facilitated the preparation and reporting of the systematic review[20]

Topic Refinement and Protocol Review

The Evidence-based Practice Centers (EPCs) developed this topic and KQs through a public process. The topic was nominated within AHRQ and subsequently developed and refined by our EPC. Initially, a panel of key informants gave input on the KQs to be examined; AHRQ then posted these questions on the Effective Health Care Website for public comment from September 15, 2014, through October 6, 2014. We revised the KQs in response to comments.

We then drafted a protocol for the systematic review and recruited a panel of technical experts to provide high-level content and methodological expertise throughout the development of the review. The final protocol was posted on the Effective Health Care website at http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2030 on December 30, 2014 and registered on PROSPERO (Registration number: CRD42015024759).

Literature Search Strategy

We systematically searched, reviewed, and analyzed the scientific evidence for each of our three KQs. We began with a focused MEDLINE® search for eligible interventions using a combination of medical subject headings (MeSH®) and title and abstract keywords, limiting the search to human-only studies (from inception through November 13, 2014). We also searched the Cochrane Library, PsycINFO, and CINAHL (Cumulative Index to Nursing and Allied Health Literature) using analogous search terms.

In addition, we searched the gray literature and manually searched the reference lists of landmark studies and background articles. Two trained research team members independently reviewed all titles and abstracts identified through searches for eligibility against our inclusion and exclusion criteria. Studies marked for possible inclusion by either reviewer underwent a dual, independent full-text review. For studies without adequate information to determine inclusion or exclusion, we retrieved the full text and then made the determination. We tracked all results in an EndNote® bibliographic database (Thomson Reuters, New York, NY).We resolved conflicts at the full-text stage by consensus.

Trained reviewers abstracted important information from included studies into evidence tables, housed on AHRQ's Systematic Review Data Repository. A second senior member of the team reviewed all data abstractions for completeness and accuracy. Reviewers resolved conflicts by discussion and consensus or by consulting a third member of the review team.

Risk of Bias Assessment

To assess the risk of bias (internal validity) of studies, two independent reviewers used predefined, design-specific criteria based on guidance in the *Methods Guide*.[21] We resolved conflicts by consensus or by consulting a third member of the team. For RCTs, we relied on the

risk of bias tool developed by the Cochrane Collaboration.[22] We assessed the risk of bias of observational studies using questions from an item bank developed by RTI International[23] and the Cochrane Risk Of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI).[24] Minimum eligibility criteria for systematic reviews included an explicit description of search strategy used and determination that the search strategy was adequate, application of predefined eligibility criteria and risk of bias assessment for all included studies, and synthesis of the results presented.

In general terms, a study with no identifiable flaws has a low risk of bias. A study with medium risk of bias is susceptible to some bias but probably not sufficient to invalidate its results. A study with high risk of bias has significant methodological flaws (stemming from, e.g., serious errors in design or analysis) that may invalidate its results. We considered the risk of bias for each relevant outcome of a study. When studies did not report sufficient detail to assess the validity of the design or study conduct, we judged the risk of bias to be unclear.

Data Synthesis

Although we had planned meta-analysis if we found five or more similar studies that use a common design (all RCTs or all cohort) for a comparison of interest,[25] we did not encounter a sufficiently large volume of similar studies and therefore relied on analyses of single studies, supplemented by reported or calculated relative risks or mean differences, with confidence intervals, whenever possible. To determine whether pooled analyses were appropriate, we assessed the clinical and methodological heterogeneity of the studies under consideration following established guidance.[26] We did this by qualitatively assessing the PICOTS of the included studies, looking for similarities and differences.

Strength of the Body of Evidence

We graded the strength of a body of evidence based on the updated guidance in the *Methods Guide*.[27, 28] The AHRQ EPC approach incorporates five key domains: study limitations (includes study design and aggregate risk of bias), consistency, directness, precision of the evidence, and reporting bias. It also considers other optional domains that may be relevant for some scenarios, such as a dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect). These domains are particularly relevant for observational studies. Thus, we considered these domains in addition to the five key domains for observational studies included in our review.

Two reviewers assessed each domain for each key outcome and resolved any differences by consensus discussion. Senior members of the review team (including at least one subject matter expert and one methodologist) graded the strength of evidence.

Grades reflect the confidence that the reviewers have that various estimates of effect are close to true effects with respect to the KQs in a systematic review. Table B defines the four grades.

Table B. Definitions of the grades of overall strength of evidence[27]

Grade	Definition
High	We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Risk of bias assessments for individual studies feed into the rating for the first of the strength of evidence domains, study limitations. Specifically, we rated bodies of evidence comprising trials with a high risk of bias as having high study limitations. Medium or unclear risk of bias studies resulted in medium study limitations. Low risk of bias studies resulted in low study limitations. In keeping with GRADE and strength of evidence guidance, we rated observational studies as having high study limitations.

As described above, study design and study limitations together set the baseline strength of evidence grade. Other domains then could either reduce or increase the grade. A body of evidence with high study limitations, with no other reasons to increase confidence (arising from dose-response, large magnitude of effect, or plausible confounding) or decrease it (arising from inconsistency, imprecision, indirectness, or reporting bias) would generally have a low strength of evidence grade. A body of evidence with low study limitations, with no reasons to decrease confidence (arising from inconsistency, imprecision, indirectness, or reporting bias), would generally have a high strength of evidence grade. In other words, although study design and study limitation provide a baseline judgment of strength of evidence, each of four additional source of uncertainty (inconsistency, imprecision, indirectness, reporting bias) serve to further reduce the strength of evidence grade.

For each source of uncertainty, we consistently used the following rubric to evaluate its effect on the overall strength of evidence across outcomes. Specifically, for indirectness, we rated intermediate outcomes as direct, rather than indirect, evidence. For this systematic review, these outcomes can be interpreted as direct measures of process change. Regarding consistency, we rated it as unknown for bodies of evidence with single studies; the rating of unknown consistency did not lower the overall grade. We relied on established guidance to judge precision.[29] Regarding imprecision, we specified the reasons for our judgment in footnotes to strength of evidence tables (small sample size or event rate, particularly when considering the optimum information size for the specific outcome, confidence intervals crossing the line of no difference or very wide confidence intervals). We downgraded the overall strength of evidence by two levels when we found multiple reasons for imprecision. We upgraded the evidence by one level for factors such as large magnitude of effect.

Applicability

We accessed applicability of the evidence following guidance from the *Methods Guide*.[55] We used the PICOTS framework to explore factors that affect applicability. Some factors relevant to the generalizability of our findings include the following:

- patient characteristics in study do not match typical characteristics of patients receiving mental health care;
- study's health care delivery setting in system or organization are not generalizable to typical settings;
- nature of the comparison usual care group is not typical of type of mental health care rendered in the system or organization or provided by practitioners;
- types of practitioners in the organization the study employed does not match those in typical mental health care settings
- the implementation of particular EBP interventions is not feasible in typical care settings;
- the intensity of the QI, implementation, or dissemination strategy employed by the study is not feasible to apply in practice;
- the timing of the strategy would be difficult to implement typical care settings

Results

We provide a summary of results by KQ below. Detailed descriptions of included studies, key points, detailed synthesis, summary tables, and expanded strength of evidence tables that include the magnitude of effect can be found in the full report. Our summary of results below presents the strength of evidence grades.

Results of Literature Searches

Figure B presents our literature search results through November 13, 2014. After applying our eligibility and exclusion criteria to titles and abstracts of all 6,086 identified citations, we obtained full-text copies of 490 articles for review. Because of the lack of standard terminology used to define the types of studies of interest to this review, we used a wide-ranging search strategy. As a result, many citations were not relevant, leading to a much smaller pool of includes at full-text. We excluded 476 of these articles at full-text, leaving 14 articles representing 15 studies[13, 14, 30-41] included in our review (one article included 2 studies within the same publication.[39]). Common reasons for exclusion included not meeting review criteria for population (n=232), comparator (n=40), intervention (n=90), publication type (n=53), outcome (n=33). We did not find any relevant non-English studies with English abstracts upon review.

This evidence base for KQ 1 consisted of 15 studies from 14 publications.[13, 14, 30-41] One of these studies addressed KQ 2 (harms) and four addressed KQ 3 (moderators of effectiveness). The evidence base included randomized controlled trials,[13, 14, 30-32, 34, 35, 37, 39-41] controlled clinical trials,[36, 38] ecological aggregate,[33] and quasi-experimental.[39] We assessed risk of bias for all 15 studies dually and independently. We assessed 6 studies as unclear risk of bias, 1 as low, 2 as medium, and 6 as high. Full evidence tables are available at http://srdr.ahrq.gov/projects/530.

We first attempted to categorize each strategy by whether it focused on implementation, dissemination, or quality improvement (QI_ according to our definitions, but we encountered several difficulties. First, our categorizations—assessed independently by two reviewers, with conflicts resolved by consensus—did not always match the study authors' categorization (dissemination, implementation or quality improvement). Second, the complexity of several of the strategies meant that we could not assign studies to mutually exclusive categories for implementation, dissemination, or quality improvement. We judged that 6 of 15 studies could be classified as having dual categories. Third, studies within the same category (dissemination, implementation, or quality improvement) did not have sufficient similarities in intervention components to enable meaningful synthesis of findings.

As a result of these difficulties, we decided to categorize each of the strategies according to the EPOC taxonomy, as described above. We classified strategies with one or more financial or organizational components as "financial or organizational change" strategies, and strategies with only professional components as "professional training" strategies. The two categories, "financial or organizational change" and "professional training," guided our qualitative synthesis. We present summary tables of descriptions of strategy components and differences by study arms for each included study in the text of our main report. Table C presents study characteristics for professional training and financial or organizational change strategies.

Figure B. Results of literature searches

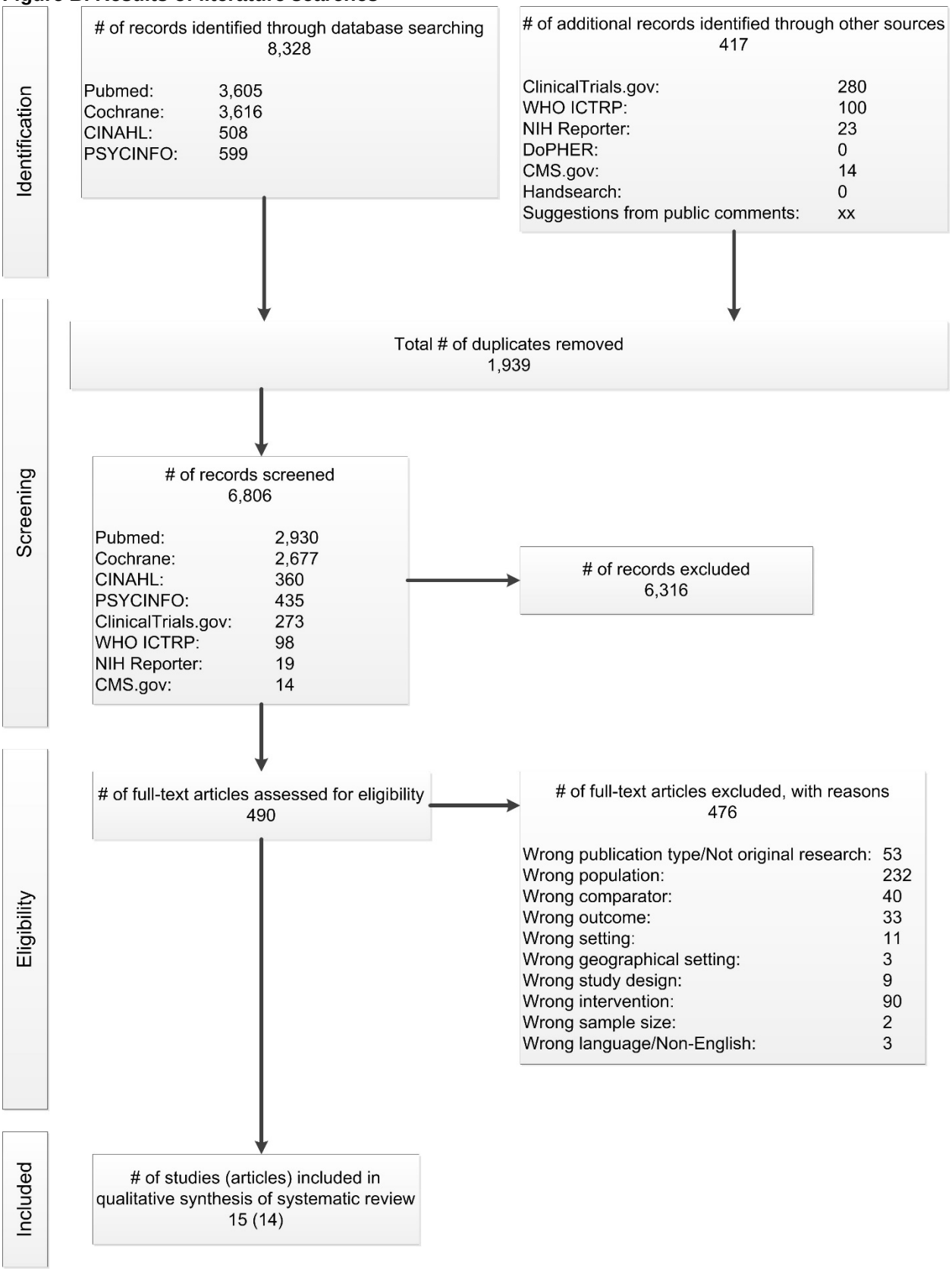


Table C. Strategies to improve mental health of children and adolescents: Study characteristics

Study Descriptor	Characteristics	Primary Strategy: Professional Training[a]	Primary Strategy: Financial or Organizational Change[b]	Total
Design	RCT	2	0	2
	2-stage RCT	0	1	1
	Cluster RCT	2	6	8
	CCT	0	2	2
	Non-RCT	2	0	2
Setting	Primary Care	1	1	2
	Community Mental Health	4	8	12
	School	1	0	1
Primary Strategy Component	Professional Training[a]	6	0	6
	Financial or Organizational Change[b]	0	9	9
Strategy Categorization[c]	Quality Improvement	2	3	5
	Implementation	1	3	4
	Dissemination	0	0	0
	Hybrid QI and I	0	2	2
	Hybrid QI and D	2	1	3
	Hybrid I and D	1	0	1
Risk of Bias	Low	1	0	1
	Medium	0	2	2
	High	3	3	6
	Unclear	2	4	6
Key Question	KQ 1	6	9	15
	KQ 2	1	0	1
	KQ 3	1	3	4
Total N of studies		6	9	15

[a] Included all professional components from the EPOC taxonomy

[b] Included at least 1 financial or organizational component from the EPOC taxonomy

[c] Categories dually assigned by members of the study team according to the definitions of QI, I, and D included in the PICOTS

ADHD = attention deficit hyperactivity disorder; CCT = controlled clinical trial; D =Dissemination; I=Implementation; KQ = Key Question; QI = quality improvement; RCT = randomized controlled trial.

Below, we summarize the main findings and strength of evidence, where applicable. We then discuss the findings in relationship to what is already known, applicability of the findings, implications for decisionmaking, limitations, research gaps, and conclusions.

Key Findings and Strength of Evidence

Key Question 1. Effectiveness of Strategies to Improve Mental Health Care for Children and Adolescents

Overview

We found the strategies included in this review to be heterogeneous and difficult to categorize. We encountered a large degree of uncertainty and overlap when classifying the examined strategies as implementation, dissemination, and quality improvement (QI) (our initial taxonomy). We then shifted to the EPOC taxonomy to identify individual components and

groups of components. This taxonomy allowed us to group strategies in two categories: (1) professional training strategies with professional components only or (2) financial or organizational change strategies, with at least one financial or organizational component in addition to professional components. Most strategies were complex and included multiple (2 to 7) different components.

Tables D and E present strength of evidence grades for professional training and financial or organizational change strategies, respectively. We graded 14 outcomes for professional training and 16 for financial or organizational change, and over half of these grades are insufficient or low for no benefit. Nonetheless, we found evidence that a majority of strategies had at least some evidence of effectiveness. Ten studies reported in 9 publications (i.e., 9 strategies) had at least one outcome rated as low for benefit, while one study had a single outcome rated as moderate for benefit. Overall, 3 of 5 strategies (6 studies) reported in 5 publications classified as having professional components only and 7 of 9 strategies classified as having at least one organizational or financial component had at least one outcome rated as low or medium for benefit. Therapists in the pay-for-performance group were over twice as likely to demonstrate implementation competence compared with implementation-as-usual therapists.[37] Other outcomes for which we found low strength of evidence of benefit included improved practitioner adherence from training practitioners to monitor metabolic markers,[33] computer decision support plus EHR that included diagnosis and treatment guidelines,[30] and Internet portal providing practitioner access to practice guidelines;[40] improved practitioner morale, engagement, and stress from organizational change;[35] improved patient access to care, parent satisfaction, treatment engagement, and therapeutic alliance from training nurses to educate parents about EBPs;[39] improved child behavior problems in the short term (6 months) and out-of-home placements from organization change,[14] improved patient functional status from weekly feedback on patient symptoms and functioning to practitioners;[13] and improved service utilization from training practitioners on medication monitoring[33] and appropriate identification and referral of cases.[32]

Only four strategies (1 study each) consistently provided insufficient or evidence of no benefit across all reported outcomes. These included a strategies testing augmented active learning versus computerized routine versus routine practitioner workshop to implement an EBP,[34] collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control,[31] an Intensive Quality Assurance system versus workshop only to implement an EBP intervention,[38] and professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control (3 arms).[41]

We found no clear patterns of effectiveness associated with categorization of the strategy according to EPOC components. The studies varied with respect to the number and types of active components (that is, differences in components that comprised the treatment group strategy and the comparison group strategy). In some studies, the treatment group contained myriad components and the comparison group contained none of those components. In other studies, both the treatment and comparison groups tested strategies with multiple components, with varying number of differences in components across arms. When both arms receive several active interventions, the Hawthorne effect may explain lack of effectiveness. We did not find, however, any consistent patterns of effectiveness regarding the number of active components, that is, we did not find that studies that employed strategies with a single active component had any better or worse effect on outcomes than those that employed multiple active components.

Table D. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: satisfaction/ acceptability 1 RCT; 115 therapists[34] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Insufficient for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, single imprecise measure, CIs cross the line of no difference)
Practitioner: adherence/ fidelity 1 RCT; 115 therapists[34] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Low for no benefit for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, multiple imprecise measures with CIs crossing the line of no difference)[a]
Practitioner: adherence/ fidelity 1 ecological study; practitioners of 2,376 patients[33] 38.3% of patients had MMT in the charts after program implementation; drop in the prevalence of SGA prescribing from 15.4% in the pre-MMTP period to 6.4% in the post-MMTP period (p<0.001)	Educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator	Low for benefit for patient medication monitoring training program for practitioners vs. usual care (high study limitations, precise outcomes)
Practitioner: competence/skills 1 RCT; 115 therapists[34] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Low for no benefit for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, multiple imprecise measures with CIs crossing the line of no difference)

Table D. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient: access to care 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[39] Improved parent ratings of access to care (mean difference between groups ranged from 0.08 to 2.1 points in Study 1 and 0.6 to 1.9 in Study 2, scale 1-5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: access to care 1 RCT; 110 practices, 79 patients[32] RR of referral to early intervention: 1.20, 95% CI, 0.74 to 1.95, p=0.48	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient: satisfaction 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[39] Improved parent ratings of satisfaction of care by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: treatment engagement 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[39] Strategy improved parent ratings of treatment engagement by a mean of 0.9 in Study 1 and 2.5 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: therapeutic alliance with provider 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[39] Strategy improved parent ratings of therapeutic alliance by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)

Table D. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health or service utilization outcome: changes in mental health status 1 RCT; 511 patients[41] CP-TF and CP-BT groups had fewer behavioral problems as rated by teachers (beta=-0.41, SE=0.16, p=0.01). There were no significant differences in teacher ratings of behavioral problem for the CP-BT versus comparison group or for any comparisons of behavioral problems as rated by parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Patient health or service utilization outcomes: changes in mental health status 1; 158 patients for detainment under Mental Health Act, 83 patients for recovery[32] No statistically significant differences between groups	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient health or service utilization outcomes: socialization skills and behaviors 1 RCT; 511 patients[41] CP-TF had fewer minor assaults than comparison group as reported by the child (beta=-0.25, SE=0.12, p=0.03) and social/academic competence as reported by the teacher (beta=0.35, SE=0.13, p=0.01). These differences were not significant for the CP-BT versus comparison groups, nor were any significant differences found between groups on social skills as rated by parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Patient health or service utilization outcomes: service utilization 1 ecological study; 253 patients (82 before, 171 after MMTP implementation (SGA-treated subset of patients from overall N of 2,376)[33]	Educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator	Low for benefit for patient medication monitoring training program for practitioners vs. usual care (high study limitations, precise results)

Table D. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health or service utilization outcomes: service utilization 1 RCT: 68 patients for number of consultations in primary care following the intervention and duration of untreated psychosis and delay in reaching early-intervention services[32] Patients in the professional training group averaged 223.8 days shorter for time from the first decision to seek care to the point of referral to an early-intervention service than patients in the control group	Educational meetings, local consensus process, educational outreach visits, marketing	Low for benefit for professional training to identify and refer cases vs. treatment as usual (high study limitations, precise results)

[a]We rated this outcome as low for no benefit rather than insufficient because of the consistency of results from multiple measures.

CI = confidence interval; CP-BT = Coping Power-basic training; CP-TF = Coping Power-training plus feedback; EBP = evidence-based practice; MMT= metabolic monitoring program; MMTP = metabolic monitoring training program; NR = not reported; RCT = randomized controlled trial; SGA = second generation antipsychotic; SE = standard error; vs. = versus.

Table E. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: adherence/fidelity 1 RCT; 84 patients[30] Practitioner adherence improved through uptake of guidelines for diagnostic assessment (aOR, 8.0; 95% CI, 1.6 to 40.6); more reporting of 3 of 4 symptom domains at diagnosis No statistically significant differences on practitioner adherence through reassessment of symptoms at 3 months, adjustment of medications, mental health referral	Patient-mediated intervention, reminders, quality monitoring	Low for benefit for computer decision support plus Electronic Health Record (EHR) that included diagnosis and treatment guidelines vs. computer decision support plus EHR only on two of three measures (uptake of guidelines and measurement of ADHD symptoms at diagnosis; medium study limitations, imprecise results with small number of events, large magnitude of effect); insufficient for reassessment of symptoms at 3 months, adjustment of medications, and referral (medium study limitations, imprecise results [CI cross the line of no difference])
Practitioner: adherence/fidelity 1 RCT; 38 practitioners[31] Interaction for uptake of titration trials β=-0.283; SE, 0.09; p<0.01 Uptake of medication monitoring trials: p=NS, details NR	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])
Practitioner: adherence/fidelity 1 RCT; 746 patients[40] Strategy appeared to improve 4 of 5 examined outcomes that measured practitioner protocol adherence/program model fidelity outcomes (range mean change in outcome between groups range from 16.6 to -50) but estimates were very imprecise, with large confidence intervals	Distribution of educational materials, educational meetings, patient-mediated interventions, audit and feedback, reminders, provider incentives, quality monitoring and staff organization	Low for benefit for internet portal providing practitioner access to practice guidelines vs. wait-list control (medium study limitations, imprecise [wide CIs])

Table E. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: adherence/fidelity 1 RCT; variable by analysis[14] No statistically significant difference between groups for caregiver-reported MST therapist, audio-coded ratings of therapist adherence, ratings of supervisor adherence	*Organizational change and an EBP intervention* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work, quality monitoring *EBP intervention only:* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, quality monitoring *Organizational change only:* Distribution of educational materials, educational meetings, educational outreach visits, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for no benefit for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, details on precision NR)
Practitioner: adherence/fidelity 1 RCT; N of practitioners unclear[13] Two-thirds of practitioners did not view Web modules	Frequency of quality monitoring mechanism	Insufficient for weekly and cumulative 90-day vs. cumulative 90-day feedback on patient symptoms and functioning to practitioners (high study limitations, unknown precision)
Practitioner: adherence/fidelity 1 CCT; 30 practitioners, N of caregiver and patient reports and monthly data points NR[38] Study does not provide sufficient detail to judge magnitude of effect	Patient incentives and quality monitoring	Insufficient for Intensive Quality Assurance system vs. workshop only to implement an EBP intervention (high study limitations, imprecise results)
Practitioner: morale, engagement, stress 1 RCT; 197 practitioners in 26 programs[35] Trends toward improving all domains, but nonoverlapping CIs for only some domains	Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for benefit for organizational change vs. control (medium study limitations, details on precision NR)

Table E. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: competence/skills 1 RCT; 38 practitioners[31] Lower odds with overlapping confidence intervals of citing obstacles in 6 of 8 measures (2 reach statistical significance)	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size]))
Practitioner: incentives 1 RCT; 49 therapists and 936 patients[37] Therapists in the P4P group were over twice as likely to demonstrate implementation competence compared with IAU therapists (Event Rate Ratio, 2.24; 95% CI, 1.12 to 4.48); patients in the P4P condition were over five times as likely to meet target implementation standards (i.e., to receive specific numbers of treatment procedures and sessions) as IAU patients (OR, 5.19; 95% CI, 1.53 to 17.62) but confidence intervals were wide	Provider incentives	Moderate for benefit for paying practitioners for performance of successful delivery of an EBP intervention vs. implementation as usual (medium study limitations, precise results)
Patient: access to care 1 CCT; 4 pediatric practices, 20,917 children with primary care visits[36] Improvement in patient access to care (attending first Triple P visit; OR, 3.10; 95% CI, 1.63 to 5.89)	Changes in scope and nature of benefits and services	Low for benefit for co-location of an EBP program in primary care vs. enhanced referral to an EBP program (high study limitations, precise results)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 144 patients[31] F score for decrease in combined parent and teacher ratings of ADHD symptoms for group x time interaction: $F_{2, 144} = 0.44$, p=0.65	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])

Table E. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 567 caregivers of youth[14] At 6 months, lower child behavior problem scores for the MST plus ARC arm, but not ARC or MST only compared with control At 18 months, no statistically significant difference in child behavior problem scores between groups 1 RCT; 615 youth[14] Lower rate of out-of-home placement for MST or ARC, but not a statistically significant added benefit of MST plus ARC compared with control (34%)	*Organizational change and an EBP intervention* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work, quality monitoring *EBP intervention only:* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, quality monitoring *Organizational change only:* Distribution of educational materials, educational meetings, educational outreach visits, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for benefit at 6 months and no benefit at 18 months for child behavior problem scores for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, precise results) Low for benefit for out-of-home placements for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, precise results)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 600[37] No statistically significant differences between groups	Provider incentives	Low for no benefit for paying practitioners for performance of successful delivery of an EBP intervention vs. implementation as usual (medium study limitations, precise results)
Patient health and service utilization outcomes: functional status 1 RCT; 340 youth,144 clinicians, 383 caregivers[13] Membership in the weekly feedback group increased the rate of decline in symptoms and functioning severity scale by 0.01	Frequency of quality monitoring mechanism	Low for benefit for weekly and cumulative 90-day vs. cumulative 90-day feedback on patient symptoms and functioning to practitioners (high study limitations, precise results)

Table E. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health and service utilization outcomes: service utilization 1; 84 patients[30] Calculated OR: 2.195; 95% CI, 0.909 to 5.303; p=0.081, reported p-value in study=0.054	Patient-mediated intervention, reminders, quality monitoring	Insufficient for computer decision support plus EHR that included diagnosis and treatment guidelines vs. computer decision support plus EHR only (medium study limitations, imprecise results [CI cross the line of no difference])

ADHD = attentional deficit hyperactivity disorder; aOR = adjusted odds ratio; ARC = Availability, Responsiveness and Continuity; CCT = controlled clinical trial; CI = confidence interval; EHR = electronic health record; EBP = evidence-based practice; MST = multisystemic therapy; NR = not reported; IAU = implementation as usual; OR = odds ratio; P4P = pay for performance; RCT = randomized controlled trial; SE = standard error; vs. = versus.

Additional heterogeneity arose from several other sources. With the exception of the two studies reported in 1 publication[39] and two studies reporting variants of a similar intervention,[14, 35] none of the other studies tested similar strategies. The outcomes of the studies varied widely, as did the settings (community-based hospitals and clinics, general practice/primary care, home-based mental health systems, schools). Differences in the target of each strategy (e.g., practitioners, practices, systems) further precluded quantitative synthesis of our findings.

The absence of evidence on several factors of interest further limits our conclusions. We found no evidence of studies examining several intermediate outcomes, particularly system-level intermediate outcomes, as well as final patient health outcomes such as co-occurring conditions, mortality, or quality of life. We also found no evidence of strategies testing several components of the EPOC taxonomy, including any regulatory components and little evidence on strategies with financial components. We rated 6 studies as having unclear risk of bias and 6 studies reported in 5 publications as having high risk of bias. Thus, out of 15 studies included in our review, only 1 study had low risk of bias and 2 had medium risk of bias. Various issues with study design, attrition, and incomplete information reported by study authors precluded most of these studies from having a low or medium risk of bias.

The uncertain or high risk of bias of most of these studies affected the overall strength of evidence grades, as did the inclusion of single studies for each strategy examined.

Key Question 2. Harms Associated With Strategies to Improve Mental Health Care for Children and Adolescents

Only one study evaluated the harms associated with strategies to improve mental health care for children and adolescents (Table F). We graded the evidence on harms as having insufficient strength, based on high study limitations and imprecise results.

Table F. Summary of evidence of harms associated with strategies to improve mental health care among children and adolescents (KQ 2)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient: adverse events 1 RCT; 110 practices, 79 patients[32] No adverse events reported.	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient: false-positive referrals 1 RCT; 110 practices, 79 patients[32] No between-group differences in false-positive referrals to primary care	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)

RCT=randomized controlled trial; vs. = versus.

Key Question 3. Moderators of the Effectiveness of Strategies to Improve Mental Health Care for Children and Adolescents

Overall, we found evidence on four strategies (from single studies) that examined moderators of the effectiveness of strategies to improve mental health care for children and adolescents (Table G). Three examined whether training intensity influenced the degree of effectiveness. We were unable to combine the findings from these studies due to the heterogeneity in the strategies being tested. We graded two of these strategies (weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only on mental health symptoms and functional status, collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control on mental health symptoms) as having insufficient strength of evidence. We graded the third strategy (paying practitioners for performance in successfully delivering of an EBP intervention vs. implementation as usual) as having low strength for benefit of the moderating effects of training intensity on both patient intermediate (access to care) and patient health and service utilization outcomes (change in mental health status). School therapists receiving more intensive training had greater improvements in patient access to care ratings (sessions scheduled) for both children and for parents and greater improvements in mental health symptoms (i.e., less externalizing behaviors) than therapists receiving less intensive training.

A fourth study examined the moderating effects of fidelity to the EBP (meeting target Adolescent Community Reinforcement Approach) used as part of the strategy. We graded the evidence on the moderating effect of fidelity on this strategy as having low strength for no benefit on patient health outcomes, and patient remission status. Also of note, we did not find studies that examined most of our a priori list of moderators such as patient characteristics, intervention characteristics other than training intensity, factors of the outer or inner setting/organizational factors, characteristics of involved individuals, process characteristics other than training fidelity, or other moderators such as length of followup.

Table G. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Patient: patient access to care 1 RCT; 110 practices, 79 patients[32] More intensive training led to improved access to care ratings (sessions scheduled) for both children and for parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Patient: treatment engagment 1 RCT; 110 practices, 79 patients[32] No significant differences between groups	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)

Table G. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3) (continued)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Practitioner: protocol adherence/program fidelity 1 RCT; 110 practices, 79 patients[32] No significant differences between groups	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; 511 patients[41] More intensive training associated with greater improvements in mental health symptoms	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; N of practitioners unclear[13] Effect sizes for child and parent ratings of symptoms improved significantly in the more intensive training group	Weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only	Insufficient for moderating effect of training intensity on weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners (high study limitations, unknown precision)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; 197 practitioners in 26 programs[35] Reduction in mental health symptoms in the compliers group was significantly greater than that seen in the control group (t(114)=-2.72, p=.008, effect size=0.25) and in the noncomplier group (t(57)= -3.568, p=.001, effect size=0.47).	Audit and feedback and clinical multidisciplinary teams	Insufficient for moderating effect of training intensity on collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])

Table G. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3) (continued)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Patient health and service utilization: functional status 1 RCT; 49 therapists and 936 patients[37] No significant moderating effect of fidelity to EBP (meeting target A-CRA) on the association between treatment group and patient remission status	Provider incentives	Low for no benefit for moderating effect of fidelity to EBPs on paying practitioners for performance in successfully delivering of an EBP intervention vs. implementation as usual (medium study limitations, precise results)
Training Intensity Patient health and service utilization: functional status 1 RCT; N of practitioners unclear[13] Effect sizes for child and parent ratings of functional status improved significantly in the more intensive training group	Weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only	Insufficient for moderating effect of training intensity on weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners (high study limitations, unknown precision)

A-CRA = Adolescent Community Reinforcement Approach; ADHD = attention deficit hyperactivity disorder; EBP = evidence-based practice; N/A = not available; RCT=randomized controlled trial; SOE =strength of evidence.

Discussion

Key Findings and Strength of Evidence

Overall, 11 studies in 10 publications of the 15 trials published in 14 articles demonstrated effectiveness as measured by benefits on at least one examined intermediate or patient health and service utilization outcome of interest. Our confidence in these results is limited the paucity of studies on any strategy. Nonetheless, moderate strength of evidence (from 1 RCT) supports pay-for-performance to improve implementation competence.[37] Low strength of evidence supports training practitioners to monitor metabolic markers and service utilization (1 observational study);[33] providing treatment guidelines through computer decision support (1 RCT)[30] or an Internet portal (1 RCT)[40] to improve practitioner adherence; changing organizational structures to improve practitioner morale, engagement, and stress (1 study)[35] and child behavior problems in the short term (6 months) and out-of-home placements (1 study);[14] training nurses to educate parents about EBPs to improve patient access to care, parent satisfaction, treatment engagement, and therapeutic alliance (1 RCT, 1 quasi-experimental study);[39] providing weekly to practitioners on patient status to improve patient functioning (1 RCT);[13] and appropriately identifying and referring cases to improve service utilization (1 RCT).[32] We were unable to judge the potential for harms associated with these strategies that may mitigate benefits based on the single included study with information on harms. In addition, the available evidence from four studies on two moderators does not permit us to make general conclusions about the conditions under which these strategies might work optimally.

Applicability

The studies in this review were focused on children with mental health and substance abuse problems. Providers of the target interventions were practitioners with professional training such as psychiatrists, psychologists, and nurses. The applicability of findings is limited to professionally trained practitioners of children and adolescents with mental health and/or substance use disorders.

This review included dissemination, implementation, and QI strategies delivered by practitioners in typical service settings. All strategies reviewed were focused at the practitioner (e.g., training practitioners) or system (e.g., implementing a new medical management system) level. Comparison conditions included usual treatment, lower-intensity versions of the strategy under study, and pre-strategy implementation cases in one study implementing a system-level strategy within a hospital.

Outcomes examined in the studies included intermediate practitioner and intermediate patient, but not intermediate system outcomes. Thus, no studies examined intermediate system outcomes such as feasibility, update, timeliness, penetration, sustainability, and resources, including costs. Several patient health outcomes of interest such as comorbidity and mortality were not examined in any included studies.

Limitations of the Systematic Review Process

Challenges in this systematic review arose with defining the intervention of interest, constructing the search strategy, and applying prespecified inclusion/exclusion criteria.The lack of consistency in the terminology used in the published literature meant that the use of self-selected descriptors such as "QI," "implementation," or "dissemination" by study authors did not conform to our a priori definitions of these types of studies or to the other similarly labelled studies in the field. Regarding searches, we ran multiple iterations over a period of 7 months. We initially mirrored the search strategy in a previously published review but had to make substantial changes to capture concepts or terms that were not indexed by the National Library of Medicine's Medical Subject Headings (MeSH).

Regarding the application of prespecified inclusion and exclusion criteria, we found that attempts to specify the population criteria to ensure greater homogeneity of included interventions resulted in challenges when reviewing the evidence. For example, our criterion that the system or clinic care for children and adolescents with *existing* mental health issues (rather than the *risk of* mental health issues only) was difficult to apply in some cases. We included a broad range of eligible comparators in our protocol (usual care, or any other QI, implementation, or dissemination strategy). In reviewing full-text studies, we encountered otherwise eligible studies in which the intervention combined both a patient-level intervention and a system-level strategy to implement or disseminate that intervention. In such cases, the use of a usual care arm did not permit the authors to draw conclusions about the effect of the implementation or dissemination strategy apart from the underlying intervention.[43-50]

Limitations of the Evidence Base

We found relatively few studies of effectiveness of strategies to improve the mental health care of children and adolescents, although there is evidence that some are effective in improving both intermediate and patient health and resource utilization outcomes. We did not find any

studies that focused on system-level intermediate outcomes, including the costs of these strategies.

The lack of a common language to describe even a basic concern such as the primary purpose of the strategies (dissemination, implementation, or QI) served as a hindrance to synthesis. Strategies varied significantly in the number of components; the reporting on these components was not always clear enough to adequately describe the strategy or fully understand the relative importance of component parts. Studies often offered limited descriptions of "usual care" arms when compared with descriptions of experimental arms. Even with limited reporting, we found wide differences in the number, intensity, and differences in services offered in "usual care" arms. These differences sharply limit our ability to make statements about the overall effectiveness of these strategies as a class.

Only one study examined harms, and three studies examined moderators of strategy effectiveness Although the field generally acknowledges the vast array of potentially influential moderators in implementation research,[51] we uncovered only four studies on two moderators (intensity and fidelity). The paucity of evidence on issues such as fidelity and adaptation further limits our understanding of the minimum change in strategy needed to achieve a significantly different process or health outcome.

We rated several outcomes as insufficient or low strength of evidence because of the underlying heterogeneity or limited number of studies on specific strategy types, system or practitioner targets, or child or adolescent conditions. In some instances, our grades were limited by high risk of bias in included studies (arising from high attrition rates, failure to adjust analyses for baseline levels of key outcomes or clustering within practitioners or clinics, and failure to account for missing data).

Our ability to derive firm conclusions on the effectiveness of included strategies was also hindered by reporting issues in the literature. Authors reported complex analyses but often did reported other issues well enough to permit an independent evaluation of the effect size,[41] precision of the effect,[30-32, 35] or risk of bias.[30, 41]

Research Recommendations

Third-party payers are paying increasing attention to quality metrics, as health care systems move to accountable care models. We found no studies on regulatory components and just one study testing the effectiveness of a financial component.[40] In addition to expanding the modest body of evidence thus far on professional training and financial or organizational change strategies, new studies should additionally evaluate regulatory and financial components to support the needs of accountable care organizations in the near future.

We did not find evidence on the majority of the outcomes that we specified a priori. Of particular note is that six strategies relied on EBPs and therefore did not report patient health outcomes.[34-36, 38, 40, 42] In instances where fidelity to the original intervention is maintained, the assumption that the same level of effectiveness will occur in a new trial is reasonable and leads to an efficient use of research funds. Not all included studies measured fidelity adequately. New strategies relying on EBPs must, at a minimum, report on fidelity so practitioners and policymakers can judge whether the strategy is, in fact, a new intervention, rather than implementation or dissemination of an existing intervention.

The risk of crossover or contamination is of particular concern in systems interventions, but only one study explicitly provided information on the risk of crossover or contamination. As noted earlier, very few studies offered information on fidelity or on unanticipated changes.

Information on pragmatic issues around implementation (fidelity, adaptation, and minimum elements necessary to achieve change) may not necessarily require new studies on strategies with existing information; support of analyses from existing studies may fill some of the gap.

Future research in this area requires appropriately timed outcome measures. One potential explanation for the lack of consistent demonstration of effectiveness across the included studies could be that studies reported on outcomes too early, before strategies had a chance to take effect. The included studies generally measured systems outcomes over the course of the intervention. One study measured adherence to CBT after 3 months of consultation,[34] and a second measured referrals to early intervention services at 4 months after intervention.[32] Although a third trial measured outcomes at 6 months from baseline, the intervention was ongoing for some portion of that period.[40] Studies generally measured patient outcomes within 6 months of completion of the trial, with two exceptions that measured outcomes at 18[14] and 24 months,[41] respectively.

The studies we found were marked by poor reporting. Concerns about the inadequacies of reporting have been noted elsewhere in the literature.[52, 53] A recent tool, the StaRI, offers standards for reporting implementation studies that, if adopted widely, can significantly improve the utility of these studies and the pace of translation of evidence into practice.[54]

Conclusions

The evidence does not permit us to have a high degree of confidence about the efficacy of any one strategy because we generally found a single study testing each strategy. We found the strongest or most consistent evidence of benefit for strategies with that provided financial benefits to practitioners to maintain fidelity to an EBP and strategies that provide professional training to improve access to EBPs.

References

1. O'Connell ME, Boat T, Warner KE, eds. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. National Academies Press: Washington, DC; 2009.

2. Perou R, Bitsko RH, Blumberg SJ, et al. Mental health surveillance among children--United States, 2005-2011. MMWR Surveill Summ. 2013 May 17;62 Suppl 2:1-35. PMID: 23677130.

3. Weisz JR, Weiss B, Han SS, et al. Effects of psychotherapy with children and adolescents revisited: a meta-analysis of treatment outcome studies. Psychol Bull. 1995 May;117(3):450-68. PMID: 7777649.

4. Dishion TJ, McCord J, Poulin F. When interventions harm. Peer groups and problem behavior. Am Psychol. 1999 Sep;54(9):755-64. PMID: 10510665.

5. Silverman WK, Hinshaw SP. The second special issue on evidence-based psychosocial treatments for children and adolescents: a 10-year update. J Clin Child Adolesc Psychol. 2008 Jan-Mar;37(1):1-7.

6. Chorpita BF, Daleiden EL, Ebesutani C, et al. Evidence-based treatments for children and adolescents: an updated review of indicators of efficacy and effectiveness. Clin Psychol-Sci Pr. 2011 Jun;18(2):154-72. PMID: WOS:000291755900009.

7. Hoagwood K, Burns BJ, Kiser L, et al. Evidence-based practice in child and adolescent mental health services. Psychiatr Serv. 2001 Sep;52(9):1179-89. PMID: 11533391.

8. Jensen PS, Goldman E, Offord D, et al. Overlooked and underserved: "action signs" for identifying children with unmet mental health needs. Pediatrics. 2011 Nov;128(5):970-9. PMID: 22025589.

9. Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: variation by ethnicity and insurance status. Am J Psychiatry. 2002 Sep;159(9):1548-55. PMID: 12202276.

10. Saloner B, Carson N, Cook BL. Episodes of mental health treatment among a nationally representative sample of children and adolescents. Med Care Res Rev. 2014 Jun;71(3):261-79. PMID: 24399817.

11. Satcher D. U.S. Public Health Service, Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Washington, DC: Department of Health and Human Services; 2008.

12. Williams J, Klinepeter K, Palmes G, et al. Diagnosis and treatment of behavioral health disorders in pediatric practice. Pediatrics. 2004 Sep;114(3):601-6. PMID: 15342827.

13. Bickman L, Kelley SD, Breda C, et al. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. Psychiatr Serv. 2011 Dec;62(12):1423-9. PMID: 22193788.

14. Glisson C, Schoenwald SK, Hemmelgarn A, et al. Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. J Consult Clin Psychol. 2010 Aug;78(4):537-50. PMID: 20658810.

15. Balas EA, Boron SA. Managing clinical knowledege for healthcare improvements. Stuttgart, Germany: Schattauer Verlagsgesellschaft; 2000.

16. Shojania KG, McDonald KM, Wachter RM, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies, Vol. 1: Series Overview and Methodology Technical Review 9 (Contract No. 290-02-0017 to the Stanford University-UCSF Evidence-based Practices Center). AHRQ Publication No. 04-0051-1. Rockville, MD: Agency for Healthcare Research and Quality; Aug 2004.

17. Institute of Medicine. Crossing the Quality Chasm. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.

18. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy. Oslo: Norwegian Knowledge Centre for the Health Services; 2002.

19. United Nations Development Program. Human development report.; 2014.

http://hdr.undp.org/en/content/human-development-report-2014. Accessed on 29 May 2015.

20. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339(4):b2535. PMID: 19622551.

21. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. AHRQ Publication No. 12-EHC047-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2012. www.effectivehealthcare.ahrq.gov/

22. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928. PMID: 22008217.

23. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012 Feb;65(2):163-78. PMID: 21959223.

24. Sterne JAC, Higgins JPT, Reeves BC, et al. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI), Version 1.0.0. 2014. http://www.riskofbias.info. Accessed on 29 May 2015.

25. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med. 2014 Feb 18;160(4):267-70. PMID: 24727843.

26. West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity Methods Research Paper. AHRQ Publication No. 10-EHC070-EF. Rockville, MD: Agency for Healthcare Research and Quality; Sep 2010. http://effectivehealthcare.ahrq.gov/

27. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290- 2007-10056-I). AHRQ Publication No. 13(14)-EHC130-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm

28. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2014 Dec 20PMID: 25721570.

29. Bhandari M, Sprague S, Hanson B, et al. Health-related quality of life following operative treatment of unstable ankle fractures: a prospective observational study. J Orthop Trauma. 2004 Jul;18(6):338-45. PMID: 15213498.

30. Carroll AE, Bauer NS, Dugan TM, et al. Use of a computerized decision aid for ADHD diagnosis: a randomized controlled trial. Pediatrics. 2013 Sep;132(3):e623-9. PMID: 23958768.

31. Epstein JN, Rabiner D, Johnson DE, et al. Improving attention-deficit/hyperactivity disorder treatment outcomes through use of a collaborative consultation treatment service by community-based pediatricians: a cluster randomized trial. Arch Pediatr Adolesc Med. 2007 Sep;161(9):835-40. PMID: 17768282.

32. Lester H, Birchwood M, Freemantle N, et al. REDIRECT: cluster randomised controlled trial of GP training in first-episode psychosis. Br J Gen Pract. 2009 Jun;59(563):e183-90. PMID: 19520016.

33. Ronsley R, Rayter M, Smith D, et al. Metabolic monitoring training program implementation in the community setting was associated with improved monitoring in second-generation antipsychotic-treated children. Can J Psychiatry. 2012 May;57(5):292-9. PMID: 22546061.

34. Beidas RS, Edmunds JM, Marcus SC, et al. Training and consultation to promote implementation of an empirically supported treatment: a randomized trial. Psychiatr Serv. 2012;63(7):660-5. PMID: 22549401.

35. Glisson C, Hemmelgarn A, Green P, et al. Randomized trial of the Availability, Responsiveness, and Continuity (ARC) organizational intervention with community-based mental health programs and clinicians serving youth. J Am Acad Child Adolesc Psychiatry. 2012 Aug;51(8):780-7. PMID: 22840549.

36. Wildman BG, Langkamp DL. Impact of location and availability of behavioral health services for

children. J Clin Psychol Med Settings. 2012 Dec;19(4):393-400. PMID: 23053830.

37. Garner BR, Godley SH, Dennis ML, et al. Using pay for performance to improve treatment implementation for adolescent substance use disorders: results from a cluster randomized trial. Arch Pediatr Adolesc Med. 2012 Oct;166(10):938-44. PMID: 22893231.

38. Henggeler SW, Sheidow AJ, Cunningham PB, et al. Promoting the implementation of an evidence-based intervention for adolescent marijuana abuse in community settings: testing the use of intensive quality assurance. J Clin Child Adolesc Psychol. 2008 Jul;37(3):682-9. PMID: 18645758.

39. Gully KJ, Price BL, Johnson MK. Increasing abused children's access to evidence-based treatment: diffusion via parents as consumers. Child Maltreat. 2008 Aug;13(3):280-8. PMID: 18359928.

40. Epstein JN, Langberg JM, Lichtenstein PK, et al. Use of an Internet portal to improve community-based pediatric ADHD care: a cluster randomized trial. Pediatrics. 2011 Nov;128(5):e1201-8. PMID: 22007005.

41. Lochman JE, Boxmeyer C, Powell N, et al. Dissemination of the Coping Power program: importance of intensity of counselor training. J Consult Clin Psychol. 2009 Jun;77(3):397-409. PMID: 19485582.

42. Chadwick Center for Children and Families. Closing the quality chasm in child abuse treatment: Identifying and disseminating best practices. San Diego, CA: Author; 2004.

43. Asarnow JR, Jaycox LH, Duan N, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. JAMA. 2005 Jan 19;293(3):311-9. PMID: 15657324.

44. Asarnow JR, Jaycox LH, Tang L, et al. Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. Am J Psychiatry. 2009 Sep;166(9):1002-10. PMID: 19651711.

45. Ngo VK, Asarnow JR, Lange J, et al. Outcomes for youths from racial-ethnic minority groups in a quality improvement intervention for depression treatment. Psychiatr Serv. 2009 Oct;60(10):1357-64. PMID: 19797376.

46. Wells KB, Tang L, Carlson GA, et al. Treatment of youth depression in primary care under usual practice conditions: observational findings from Youth Partners in Care. J Child Adolesc Psychopharmacol. 2012 Feb;22(1):80-90. PMID: 22251025.

47. Henggeler SW, McCart MR, Cunningham PB, et al. Enhancing the effectiveness of juvenile drug courts by integrating evidence-based practices. J Consult Clin Psychol. 2012 Apr;80(2):264-75. PMID: 22309470.

48. Liddle HA, Dakof GA, Henderson C, et al. Implementation outcomes of Multidimensional Family Therapy-Detention to Community: a reintegration program for drug-using juvenile detainees. Int J Offender Ther Comp Criminol. 2011 Jun;55(4):587-604. PMID: 20427547.

49. Weisz JR, Southam-Gerow MA, Gordis EB, et al. Cognitive-behavioral therapy versus usual clinical care for youth depression: an initial test of transportability to community clinics and clinicians. J Consult Clin Psychol. 2009 Jun;77(3):383-96. PMID: 19485581.

50. Bryant J, Passey ME, Hall AE, et al. A systematic review of the quality of reporting in published smoking cessation trials for pregnant women: an explanation for the evidence-practice gap? Implement Sci. 2014;9:94. PMID: 25138616.

51. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:50. PMID: 19664226.

52. Rycroft-Malone J, Burton CR. Is it time for standards for reporting on research about implementation? Worldviews Evid Based Nurs. 2011;8:189-90.

53. Newhouse R, Bobay K, Dykes PC, et al. Methodology issues in implementation science. Med Care Res Rev. 2013;51(Suppl 2):S32-40.

54. Pinnock H, Epiphaniou E, Sheikh A, et al. Developing standards for reporting implementation studies of complex interventions (StaRI): a

systematic review and e-Delphi. Implement Sci. 2015;10(1):42. PMID: 25888928.

Introduction

Background

Condition

Approximately one in five children and adolescents living in the United States has one or more mental, emotional, or behavioral health disorders according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) criteria in any given year.[1] These disorders contribute to problems with family, peers, and academic functioning. They may exacerbate coexisting conditions (including other mental and substance use disorders and chronic health conditions) and may reduce quality of life. They also increase the risk of involvement with the criminal justice system and other risk-taking behaviors and suicide.[2]

Strategies to Improve Mental Health in Children

Several key publications in the mid- to late-1990s suggested that usual care in children's mental health had, at best, no[3] and sometimes harmful effects.[4] Since then, the evidence base for pediatric mental health interventions that target mood disorders, anxiety disorders, disruptive behavior disorders, psychotic disorders, eating disorders, and substance use disorders has grown.[5, 6]

Despite advances in the evidence base,[5, 7] outcomes for children with mental health problems remain suboptimal because of issues with access to care and the failure of systems and providers to adopt established quality improvement (QI) strategies and interventions with proven effectiveness (e.g., evidence-based practices [EBPs]). Studies using nationally representative data on U.S. adolescents show that only approximately one in five children with mental health problems receives services, and only one-third of treatment episodes are considered minimally adequate (at least four visits with psychotropic medication or at least eight visits without psychotropic medication).[8-10] The current health care system continues to provide fragmented care to children and adolescents in numerous uncoordinated systems, rendering inefficient the delivery of needed services.[11] Moreover, clinicians (particularly primary care practitioners) may lack the time, knowledge, or training to identify and treat or refer mental health problems appropriately.[12]

Given the gap between observed and achievable processes and outcomes, the next step is the adoption of QI strategies and the development of strategies to implement or disseminate these interventions. Such strategies target changes in the organization and delivery of mental health services.[13, 14] They seek to improve the quality of care and patient outcomes by closing the gap between research evidence and practice.[15-17]

In keeping with recent AHRQ reviews with a similar focus,[18] we view QI strategies as "any intervention aimed at reducing the quality gap for a group of patients representative of those encountered in routine practice."[16, p.13] For this review, we focus on QI strategies targeting practitioners (e.g., via education, training, and supervision) and organizations (e.g., via financial incentives, regulation, and policies) that provide mental health care to children and adolescents, with the ultimate goal of improving both the process and the outcomes of that care.[19, 20]

Some investigators consider dissemination and implementation strategies as a particular subset of initiatives to improve the quality of care. However, the field of implementation and dissemination is so new that the conceptual framework and terminology in relationship to QI

efforts have not yet been fully standardized.[21] We do not take a position on the taxonomy of these terms but refer in the remainder of this report to these strategies as QI, implementation, and dissemination.

Dissemination strategies can be differentiated from implementation strategies.[22-25] Dissemination is "the active and targeted distribution of information and interventions to a specific public health or clinical practice audience via determined channels using planned strategies," with "the intent to spread knowledge and associated evidence-based interventions to enhance the adoption and the implementation of the information or intervention."[22, p.2] Implementation is "the use of strategies to integrate evidence-based health interventions (e.g., EBPs) and change practice patterns within specific settings."[22, p.2]

The taxonomy used by the international Cochrane Review Group's Effective Practice and Organisation of Care (EPOC) Group, which studies these types of often times complex strategies designed to improve health care professionals' practice and the organization of health care services classifies these strategies as having one or more professional, financial, organizational, and regulatory components. These strategies that ultimately strive to improve practice and organization of services typically include various forms of continuing education for providers; quality assurance projects; and financial, organizational, or regulatory interventions that can affect the ability of health care professionals to deliver services more effectively and efficiently.

The ultimate goal of these strategies is to improve patient health and service utilization outcomes for children and adolescents with mental health problems. Intermediate outcomes in this context include changes to health care systems, organizations, and practitioners that provide mental health care. Targeting multiple, interrelated, nested levels such as the macro environment (e.g., state), organization or system (e.g., specialty mental health clinic), program (e.g., selected intervention), practitioners (e.g., clinicians), and patients (e.g., children or adolescents and their families) typically increase the effectiveness and sustainability of a particular strategy.[26, 27] For instance, changes in intermediate outcomes such as practitioners' attitudes[28] or organizational climate[29] may influence the successful adoption of and fidelity to EBPs. These practices in turn influence patient health outcomes, such as behavior or quality of life.

Potential Moderators of Strategy Effectiveness

Several frameworks guide investigations of how certain variables, including contextual factors, influence the effectiveness of the QI, implementation, or dissemination strategy.[30-34] For example, factors such as the diversity of outpatient settings, which may include schools, primary care, specialty mental health, emergency rooms, and, increasingly, homes for children's mental health services, may influence the generalizability and applicability of QI, implementation, or dissemination efforts. The organizational factors of the clinical setting may influence outcomes, and many have argued that these unique factors should be examined within the context of QI, implementation, and dissemination studies.[35, 36]

One framework commonly used to study implementation research, the Consolidated Framework for Implementation Research (CFIR),[32] comprises five major domains:

1. intervention characteristics, strength of the evidence base behind the intervention;
2. "inner setting", e.g., culture, leadership, and engagement of health care organizations);
3. "outer setting", e.g., patient needs and resources, external policies and incentives);
4. characteristics of involved individuals, e.g., professional training, experience or characteristics of parents/caregivers; and
5. process by which implementation is accomplished, e.g., plan, evaluate, and reflect.

This CFIR framework can be applied to research on effective implementation of mental health strategies for children and adolescents to begin to understand salient contextual factors.[37] We used the CFIR as an organizing framework for moderators of strategy effectiveness. In addition to the five domains of the CFIR described above, we added another category, namely, characteristics of the patient.

Scope and Key Questions

Rationale for Evidence Review

The increasing interest in strategies to improve professional practice and delivery of effective mental health services for children and adolescents with mental health problems indicates that the existing body of evidence on these strategies stands poised for an objective systematic review. Decisionmakers are in critical need of information about these approaches to improve children's mental health care. A better understanding of the comparative benefits, harms, and modifiers of the available strategies to achieve these improvements may help guide a wide array of interests, particularly for practitioners and administrators of care facilities, organizations, and health systems. Such information will also prove important for those making insurance coverage and other policy decisions for these patients with mental health care needs.

This review focuses on evidence about strategies that aim to improve the quality of mental health care rather than evidence about the efficacy or effectiveness of specific interventions. We concentrate on efforts that target practitioners or organizations/systems that care for children and adolescents with mental health problems.

Proposed Contributions to the Evidence Base

Two recent systematic reviews have addressed this topic. In 2012, Barwick and colleagues examined knowledge translation interventions and strategies related to the delivery, organization, or receipt of child and youth mental health services.[38] Most focused on practitioner or teacher training for behavior change. This systematic review excluded studies of children with substance abuse. In 2013, Novins and colleagues focused on the dissemination and implementation of mental health EBPs, including substance abuse, for children and adolescents.[39]

Scope of the Review

As reflected in our Key Questions (KQs) and analytic framework below, we have three primary aims for this review. First, we will increase knowledge about the effectiveness of dissemination, implementation, or QI strategies that seek to improve the mental health care of children and adolescents. Second, we will examine harms associated with these strategies. Third, we will attempt to determine whether effectiveness or harms vary in subgroups based on system, organizational, practitioner, or patient characteristics.

Based on feedback from our key informants, we did not attempt to review studies that focused on strategies that target systems, organizations, or providers who treat children and adolescents who have only developmental disorders, owing to heterogeneity in strategies used and types of systems involved in their care.

Ultimately, this review will inform mental health clinicians, health care system and organization administrators, policymakers, and researchers about effective ways to improve care to children and adolescent with mental health problems.

Key Questions

KQ 1: What is the effectiveness of quality improvement, implementation, and dissemination strategies employed in outpatient settings by health care practitioners, organizations, or systems that care for children and adolescents with mental health problems to improve:

a. intermediate patient, provider, or system outcomes
b. patient health and service utilization outcomes?[1]

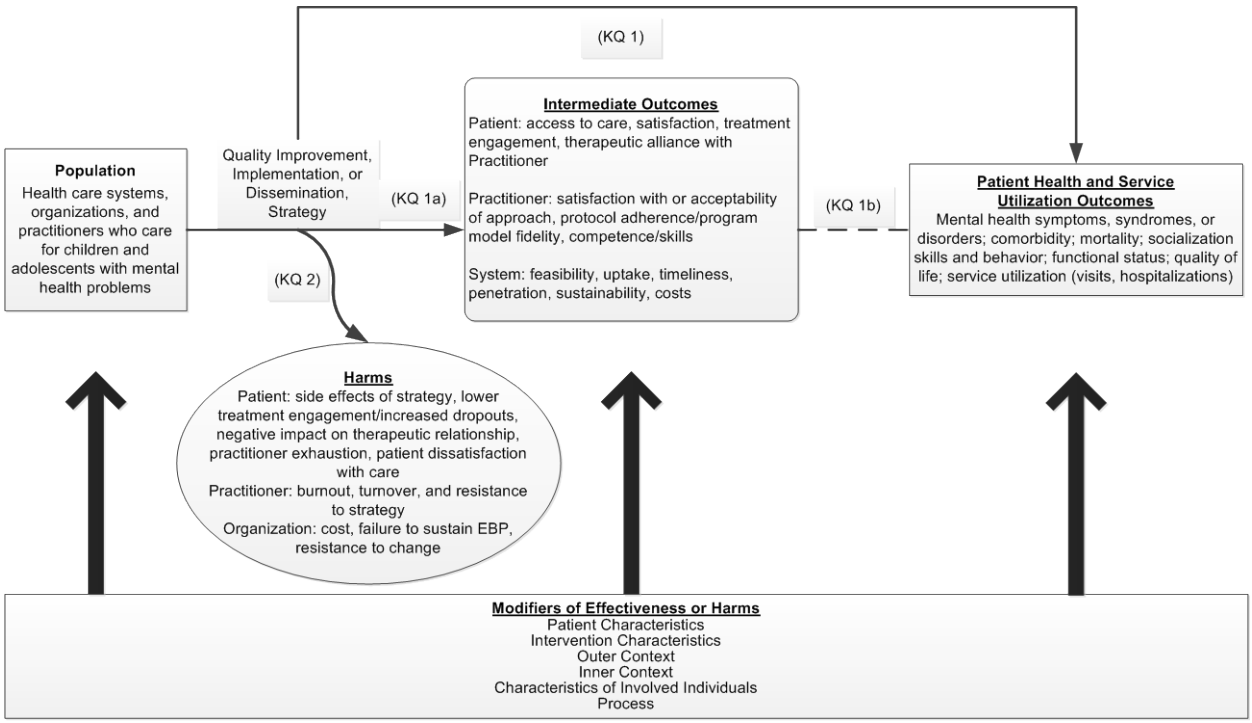
KQ 2: What are the harms of these mental health strategies?

KQ 3: Do characteristics of the child or adolescent or contextual factors (e.g., characteristics of patients, practitioners, organizations, or systems; intervention characteristics; setting; or process) modify the effectiveness or harms of strategies to improve mental health care and, if so, how?

Analytic Framework

Figure 1 depicts the patient populations, interventions, comparators, outcomes, and timing of outcomes assessment (PICOTs) and key questions in relation to these PICOTs.

Figure 1. Analytic framework for strategies to improve mental health care in children and adolescents



[1] We revised KQ 1 and the outcome specified in our protocol slightly for clarity. We replaced the term "health care providers" with "health care practitioners" to indicate that this particular phrase refers to individuals rather than systems or institutions. We also replaced "final outcomes" with "patient health and service utilization outcomes" for clarity.

Organization of This Report

We describe our methods (Chapter 2) and present our key findings in Chapter 3. In Chapter 4, we discuss our findings; we also examine the limitations of the evidence base and this review, clarify gaps in the knowledge base, and offer recommendations for future research. References follow the final chapter.

The main report has several appendixes, as follows: A, search strategies; B, excluded studies; C, risk of bias tables; D, forest plots; and, E, EPOC taxonomy tables. Evidence tables can be accessed at http://srdr.ahrq.gov/projects/530.

Methods

The methods for this systematic review follow the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* from the Agency for Healthcare Research and Quality (AHRQ; available at http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm). The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist facilitated the preparation and reporting of the systematic review.[40]

Topic Refinement and Protocol Review

The Evidence-Based Practice Centers (EPC) developed this topic and key questions (KQs) through a public process. The topic was nominated within AHRQ and subsequently developed and refined by our EPC. Initially, a panel of key informants gave input on the KQs to be examined; AHRQ then posted these questions on the Effective Health Care (EHC) Website for public comment from September 15, 2014 through October 6, 2014. We revised the KQs in response to comments.

We then drafted a protocol for the systematic review and recruited a panel of technical experts to provide high-level content and methodological expertise throughout the development of the review. The final protocol was posted on the EHC website at http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2030 on December 30, 2014 and registered on PROSPERO (Registration number: CRD42015024759).

Literature Search Strategy

Search Strategy

We systematically searched, reviewed, and analyzed the scientific evidence for each of our three KQs. We began with a focused MEDLINE® search for eligible interventions using a combination of medical subject headings (MeSH®) and title and abstract keywords, limiting the search to human-only studies (Appendix A) (from inception through November 13, 2014). We also searched the Cochrane Library, PsycINFO, and CINAHL (Cumulative Index to Nursing and Allied Health Literature) using analogous search terms. These searches included randomized controlled trials (RCTs), controlled clinical trials (CCTs), and systematic reviews. We selected these databases based on preliminary searches and consultation with content experts. We conducted quality checks to ensure that the search identified known studies (e.g.., studies identified during topic nomination and refinement). If we did not identify the known studies, we revised and reran our searches.

In addition, we searched the gray literature for unpublished studies relevant to this review and included studies that met all the inclusion criteria and contain enough methodological information to assess risk of bias. Sources of gray literature include ClinicalTrials.gov, the World Health Organization's International Clinical Trials Registry Platform, the National Institutes of Health Research Portfolio Online Reporting Tools, the Database of Promoting Health Effectiveness Reviews, and CMS.gov. To avoid retrieval bias, we manually searched the reference lists of landmark studies and background articles on this topic to look for any relevant citations that our electronic searches might have missed.

Inclusion and Exclusion Criteria

We specified our inclusion and exclusion criteria based on the populations, interventions, comparators, outcomes, timing, and settings (PICOTS) identified through the topic refinement exercise (Table 1).

Table 1. Inclusion/exclusion criteria for strategies to improve mental health services for children and adolescents

Category	Inclusion	Exclusion
Population	Health care systems, organizations, and practitioners that care for children and adolescents or mixed (child and adult) populations with mental health problems	Health care systems, organizations, and practitioners that care only for adults 18 years of age or older Health care systems, organizations, and practitioners that care for children and adolescents with only developmental disorders
Interventions (Strategies)	• Quality improvement strategies (e.g., strategies targeting systems and practitioners of mental health care to children and adolescents with the goal of improved quality of care) • Implementation strategies (e.g., strategies to integrate evidence-based practice (EBP) interventions that meet NREPP inclusion criteria with the goal of changing practice patterns) • Dissemination strategies (e.g., strategies to enhance the adoption and the implementation of evidence-based interventions that meet NREPP inclusion criteria)	Interventions targeting only patients, only drug interventions, and interventions not otherwise described in inclusion criteria
Comparator	Any control strategy, including usual care or different variants of the same intervention	None
Outcomes	Intermediate outcomes (at least one intermediate outcome is required for KQs 1, 3) • Patient – access to care – satisfaction – treatment engagement – therapeutic alliance with practitioner • Practitioner – satisfaction with or acceptability of approach – protocol adherence/program model fidelity – competence or skills • System or organization – feasibility – uptake – timeliness – penetration – sustainability – resources (including costs)	All outcomes not otherwise specified

Table 1. Inclusion/exclusion criteria for strategies to improve mental health services for children and adolescents (continued)

Category	Inclusion	Exclusion
Outcomes (continued)	Patient health and service utilization outcomes (at least one of these outcomes is required for KQs 1 and 3 unless the strategy uses an intervention that is an EBP) • Change in mental health status, including symptom change, response, remission, relapse, and recurrence • Coexisting physical health conditions, substance use problems, developmental disorders, other mental health problems • Mortality • Socialization skills and behavior • Functional status • Quality of life Service utilization (e.g., visits, hospitalizations) Harms of strategy • Patient – lower treatment engagement or more dropouts – negative impact on therapeutic relationship – side effects of evidence-based practice incorporated into strategy (e.g., adverse events, suicidality) – patient dissatisfaction with care • Practitioner – burnout or exhaustion – turnover – resistance to the intervention • System or organization – cost – failure to sustain the evidence-based practice – resistance to change	
Timing of outcome measurement	All	None
Settings	Outpatient settings serving children and adolescents with mental health problems (primary care, specialty care, emergency rooms, community mental health centers, integrated care settings, federally qualified health centers, schools, homes)	Inpatient or residential treatment settings, drug treatment programs, jails or prisons
Geographic setting	Countries with a very high human development index (HDI)[41]	Countries with high, medium, low, or very low HDI
Publication language	English	All other languages

Table 1. Inclusion/exclusion criteria for strategies to improve mental health services for children and adolescents (continued)

Category	Inclusion	Exclusion
Study design	KQs 1, 3 (benefits) • RCTs • CCTs • Systematic review and meta-analyses • Cohort studies • Interrupted time series • Case-control studies KQs 2, 3 (harms): • RCTs • CCTs • Systematic review and meta-analyses • Cohort studies • Interrupted time series • Case-control studies	• Case series • Case reports • Nonsystematic reviews • Cross-sectional studies • Before and after studies without time series data • Other designs without a control or comparison group
Publication type	Any publication reporting primary data	• Publications not reporting primary data

CCT = controlled clinical trial; EBP = evidence-based practice; D = dissemination; HDI = Human Development Index; I = implementation; KQ = Key Question; NREPP = National Registry of Evidence-based Programs and Practices; QI = quality improvement; RCT = randomized controlled trial.

We included quality improvement (QI), implementation, and dissemination strategies that targeted systems, organizations, or practitioners of mental health care to children and adolescents ages 18 years of age or younger, who were already experiencing mental health symptoms. We did not include strategies such as the implementation of educational interventions for reading disorders. We also limited our review of implementation strategies to those focusing on evidence-based practice (EBP) interventions. For defining EBPs, we relied on the minimum requirements set forth by the Substance Abuse and Mental Health Services Administration's (SAMHSA's) National Registry of Evidence-based Programs and Practices (NREPP) (www.nrepp.samhsa.gov). These criteria specify that the intervention needs to have produced one or more positive behavioral outcomes in at least one study using an experimental or quasi-experimental design with results published in a peer-reviewed journal or similar publication. In addition, implementation materials, training and support resources, and quality assurance procedures for these interventions need to be ready for use by the public.

We use the term "strategy" to reference the total sum of components used to target health care systems and/or practitioners to improve the quality of care for children and adolescents with mental health problems. We use the term "intervention" to denote a specific EBP used as part of a strategy. Examples of types of included strategies are outlined below.

1. QI: strategies targeting systems and providers of mental health care to children and adolescents with the goal of improved quality of care. Examples of QI strategies include the following:[16, 18]
   a. organization or system targets: changes to the organization including case management, changing from paper to computer systems, increased staffing, changes in reimbursement schemes;
   b. clinician targets: audit and feedback, facilitated relay of clinical data to providers, pay for performance, and provider reminder systems.
2. Implementation: strategies used to adopt and integrate EBPs (defined based on the minimum criteria set forth by SAMHSA's NREPP) into routine care (e.g., strategies to

integrate evidence-based health interventions and change practice patterns). Examples of implementation strategies that vary by method of implementation facilitation include:
   a. planning
   b. educating
   c. financing
   d. restructuring
   e. managing quality, and
   f. attending to policy contexts.[42]
3. Dissemination: strategies used to disseminate evidence through increasing access to EBPs, people's motivation to use and apply EBPs (defined based on the minimum criteria set forth by the U.S. Substance Abuse and Mental Health Services Administration in its National Registry of Evidence-based Programs and Practices [NREPP]), or people's ability to use and apply EBPs. Examples of such approaches include the following:
   a. increasing the reach of the evidence (e.g., social media, interpersonal outreach);
   b. increasing people's motivation to use and apply the evidence (e.g., use of opinion leaders, champions, social networks);
   c. increasing people's ability to use and apply the evidence (e.g., additional resources, skills building); and
   d. using a multipronged approach with any of these three dissemination strategies (e.g., social marketing, academic detailing).[22]

Because strategies tended to be complex in nature and the number and types of components that varied between the treatment arm and comparison group arm differed by study, we also recorded components of each strategy by study arm according to the EPOC taxonomy.[43] Because many of the comparison groups also contained several components, we marked the components contained in each study arm of each study. This allowed us to fully describe the numerous components that were being combined and tested in each strategy, as well as enabled us to determine whether the study arms differed by a single or multiple components.

We required each included study to report at least one intermediate outcome in a minimum of one of three major categories: (1) practitioner intermediate outcomes (satisfaction, adherence, fidelity, competence), (2) system intermediate outcomes (feasibility, uptake, timeliness, penetration, sustainability, costs), and (3) patient intermediate outcomes (access to care, satisfaction, engagement, therapeutic alliance). This helped to ensure that each included study demonstrated impact based on its stated goals of improving quality or implementing or disseminating evidence-based interventions. We also required each study to report at least one patient health or service utilization outcome (change in mental health status, comorbid conditions, mortality, socialization skills and behavior, functional status, quality of life, service utilization) if the strategy was not implementing or disseminating an EBP intervention (i.e., an intervention with proven effectiveness).

For all KQs, we excluded study designs without comparison groups to ensure that our pool of included studies provided evidence on the causal link between the strategy and outcomes. We also required that the comparator enabled examination of the strategy effectiveness. That is, we excluded studies in which the strategy (system, organizational, practitioner targets) and the intervention being tested both differed between groups, because the effectiveness of the QI, implementation, or dissemination strategy could not be isolated from the baseline intervention effects.

For KQ 1 studies of benefits and KQ 3 studies of moderators of benefits, we had planned to limit our evidence base to randomized controlled trials (RCTs) (standard, clustered, stepped-wedge), controlled clinical trials (CCTs, not randomized), systematic reviews, or meta-analyses. We also planned to consider other designs—specifically, cohort studies (prospective, retrospective, and historical control), interrupted time-series, and case-control studies that met all other inclusion and exclusion criteria—if we found sparse evidence to answer these KQs using trials and systematic reviews (with or without meta-analyses). For KQ 2 and KQ 3 studies of moderators of harms, we included experimental studies noted above, interrupted time-series, and observational evidence from prospective cohort studies, retrospective cohort studies, and case-control studies that met all other inclusion and exclusion criteria.

Our exclusion of non–English-language studies is based on limitations of time and resources. However, we examined English language abstracts of non–English-language studies to assess the potential size of the literature that would be missed through this approach.

Moderators

We searched for studies with information on the following seven moderators of effectiveness or harms. Categories 2 through 6 are consistent with the Consolidated Framework for Implementation Research (CFIR) framework defined earlier.

1. Patient characteristics, such as age, gender, race and ethnicity, cognitive ability, diagnosis and severity of mental health problem, coexisting conditions, and cotreatments;
2. Intervention characteristics, such as complexity; manualized or not; intensity, frequency or duration; and adjustment of intervention to fit context;
3. Outer setting, such as external policy, incentives, availability of alternative care systems;
4. Inner setting or organizational factors, such as type of outpatient setting, structure or size, culture, implementation climate, and readiness of organization for implementation;
5. Characteristics of involved individuals such as type, knowledge, beliefs, self-efficacy, leadership, education, certifications, and years of practice of practitioners or characteristics of parents/caregivers;
6. Process characteristics, such as fidelity to the planned strategy, fidelity to the EBP, use of champions, and supervision or oversight; and
7. Other components, such as length of followup.

Study Selection

Two trained research team members independently reviewed all titles and abstracts identified through searches for eligibility against our inclusion and exclusion criteria. Studies marked for possible inclusion by either reviewer underwent a dual, independent full-text review. For studies without adequate information to determine inclusion or exclusion, we retrieved the full text and then made the determination. We tracked all results in an EndNote® bibliographic database (Thomson Reuters, New York, NY).

We retrieved and reviewed the full text of all articles included during the title and abstract review phase. Two trained team members independently reviewed each full-text article for inclusion or exclusion based on the eligibility criteria described above. If both reviewers agreed that a study did not meet the eligibility criteria, we excluded the study. If the reviewers disagreed, conflicts were resolved by discussion and consensus or by consulting a third member of the review team. All results were tracked in an EndNote database. We also recorded the main

reason that each excluded full-text publication did not satisfy the eligibility criteria (Appendix B).

Data Extraction

For studies that met our inclusion criteria, trained reviewers abstracted important information into evidence tables. We designed data abstraction forms (in AHRQ's Systematic Review Data Repository) to gather pertinent information from each article. Data recorded included the strategies (including evidence-based interventions), characteristics of the target(s) of the specific strategy (such as systems, organizations, and clinicians), comparators, settings, characteristics of the children or adolescents with mental health problems, study designs, analysis methods, and results. A second member of the team reviewed all data abstractions for completeness and accuracy.

For systematic reviews with or without meta-analyses), we planned to use the five-step process described in the AHRQ *Methods Guide*[44] to assess the relevance and quality of the systematic review and to determine how to use the information provided. We intended then either to incorporate existing systematic reviews into this one or to use them to replace all or part of the de novo process or refine our search strategy only if they were fully relevant and of high quality. Reviews that did not meet these criteria would be used to cross-check references.

Risk of Bias Assessment

To assess the risk of bias (internal validity) of studies, two independent reviewers used predefined, design-specific criteria based on guidance in the *Methods Guide* (Appendix C).[45] We resolved conflicts by consensus or by consulting a third member of the team. For RCTs, we relied on the risk of bias tool developed by the Cochrane Collaboration.[46] We assessed the risk of bias of observational studies using questions from an item bank developed by RTI International[47] and the Cochrane Risk Of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI).[48] Minimum eligibility criteria for systematic reviews included an explicit description of search strategy used and determination that the search strategy was adequate, application of predefined eligibility criteria and risk of bias assessment for all included studies, and synthesis of the results presented.

In general terms, a study with no identifiable flaws has a low risk of bias. A study with medium risk of bias is susceptible to some bias but probably not sufficient to invalidate its results. A study with high risk of bias has significant methodological flaws (stemming from, e.g., serious errors in design or analysis) that may invalidate its results. We considered the risk of bias for each relevant outcome of a study. When studies did not report sufficient detail to assess the validity of the design or study conduct, we judged the risk of bias to be unclear.

Data Synthesis

We had planned that if we found five or more similar studies that use a common design (all RCTs or all cohort) for a comparison of interest, we would consider quantitative analysis (i.e., meta-analysis) of the data from those studies.[49] We also planned to consider conducting mixed treatment comparisons meta-analysis using Bayesian methods to compare interventions with one another if we were able to identify a sufficient number of studies with a common comparator (e.g., waitlist). For all analyses, we intended to use random-effects models to estimate pooled or comparative effects if quantitative analyses were warranted. For all outcomes, we presented

relative risks or mean differences, with confidence intervals, whenever calculable. For outcomes with multiple measures, we present forest plots (Appendix D).

To determine whether quantitative analyses were appropriate, we assessed the clinical and methodological heterogeneity of the studies under consideration following established guidance.[50] We did this by qualitatively assessing the PICOTS of the included studies, looking for similarities and differences.

Strength of the Body of Evidence

We graded the strength of a body of evidence based on the updated guidance in the *Methods Guide.*[51, 52] The AHRQ EPC approach incorporates five key domains: study limitations (includes study design and aggregate risk of bias), consistency, directness, precision of the evidence, and reporting bias. It also considers other optional domains that may be relevant for some scenarios, such as a dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect). These domains are particularly relevant for observational studies. Thus, we considered these domains in addition to the five key domains for observational studies included in our review.

Two reviewers assessed each domain for each key outcome and resolved any differences by consensus discussion. Senior members of the review team (including at least one subject matter expert and one methodologist) graded the strength of evidence.

Grades reflect the confidence that the reviewers have that various estimates of effect are close to true effects with respect to the KQs in a systematic review. Table 2 defines the four grades.

Table 2. Definitions of the grades of overall strength of evidence[51]

Grade	Definition
High	We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Risk of bias assessments for individual studies feed into the rating for the first of the strength of evidence domains, study limitations. Specifically, we rated bodies of evidence comprising trials with a high risk of bias as having high study limitations. Medium or unclear risk of bias studies resulted in medium study limitations. Low risk of bias studies resulted in low study limitations. In keeping with GRADE and strength of evidence guidance, we rated observational studies as having high study limitations.[52, 53]

As described above, study design and study limitations together set the baseline strength of evidence grade. Other domains then could either reduce or increase the grade. A body of evidence with high study limitations, with no other reasons to increase confidence (dose-

response, large magnitude of effect, plausible confounding) or decrease it (inconsistency, imprecision, indirectness, reporting bias) would generally have a low strength of evidence grade. A body of evidence with low study limitations, with no reasons to decrease confidence (inconsistency, imprecision, indirectness, reporting bias), would generally have a high strength of evidence grade. In other words, although study design and study limitation provide a baseline judgment of strength of evidence, each of four additional source of uncertainty (inconsistency, imprecision, indirectness, reporting bias) serve to further reduce the strength of evidence grade.

For each source of uncertainty, we consistently used the following rubric to evaluate its effect on the overall strength of evidence across outcomes. Specifically, for indirectness, we rated intermediate outcomes as direct, rather than indirect evidence. For this systematic review, these outcomes can be interpreted as direct measures of process change. Regarding consistency, we rated it as unknown for bodies of evidence with single studies; the rating of unknown consistency did not lower the overall grade. We relied on established guidance to judge precision.[54] Regarding imprecision, we specified the reasons for our judgment in footnotes to strength of evidence tables (small sample size or event rate, particularly when considering the optimum information size for the specific outcome, confidence intervals [CIs] crossing the line of no difference or very wide CIs). We downgraded the overall strength of evidence by two levels when we found multiple reasons for imprecision. We upgraded the evidence by one level for factors such as large magnitude of effect.

Applicability

We accessed applicability of the evidence following guidance from the *Methods Guide*.[55] We used the PICOTS framework to explore factors that affect applicability. Some factors relevant to the generalizability of our findings include the following:

- patient characteristics in study do not match typical characteristics of patients receiving mental health care;
- study's health care delivery setting in system or organization are not generalizable to typical settings;
- nature of the comparison usual care group is not typical of type of mental health care rendered in the system or organization or provided by practitioners;
- types of practitioners in the organization the study employed does not match those in typical mental health care settings
- the implementation of particular EBP interventions is not feasible in typical care settings;
- the intensity of the QI, implementation, or dissemination strategy employed by the study is not feasible to apply in practice;
- the timing of the strategy would be difficult to implement typical care settings

Peer Review and Public Commentary

Experts in dissemination, implementation, and QI strategies to improve the mental health care of children and adolescents were invited to provide external peer review of the draft systematic review. AHRQ staff and an Associate Editor reviewed the draft systematic review before it went out for peer review. The EPC Associate Editors are leaders in their respective fields and are actively involved as directors or leaders at their EPCs. Their role is to assess adherence to established methodology and guidelines for EPC-based research. The draft report

will be posted on the AHRQ Web site for 4 weeks to elicit public comment. We will respond to all reviewer comments and note any resulting revisions to the text in the "Disposition of Comments Report." This disposition report will be made available 3 months after the final systematic review is posted on the AHRQ Web site.

Results

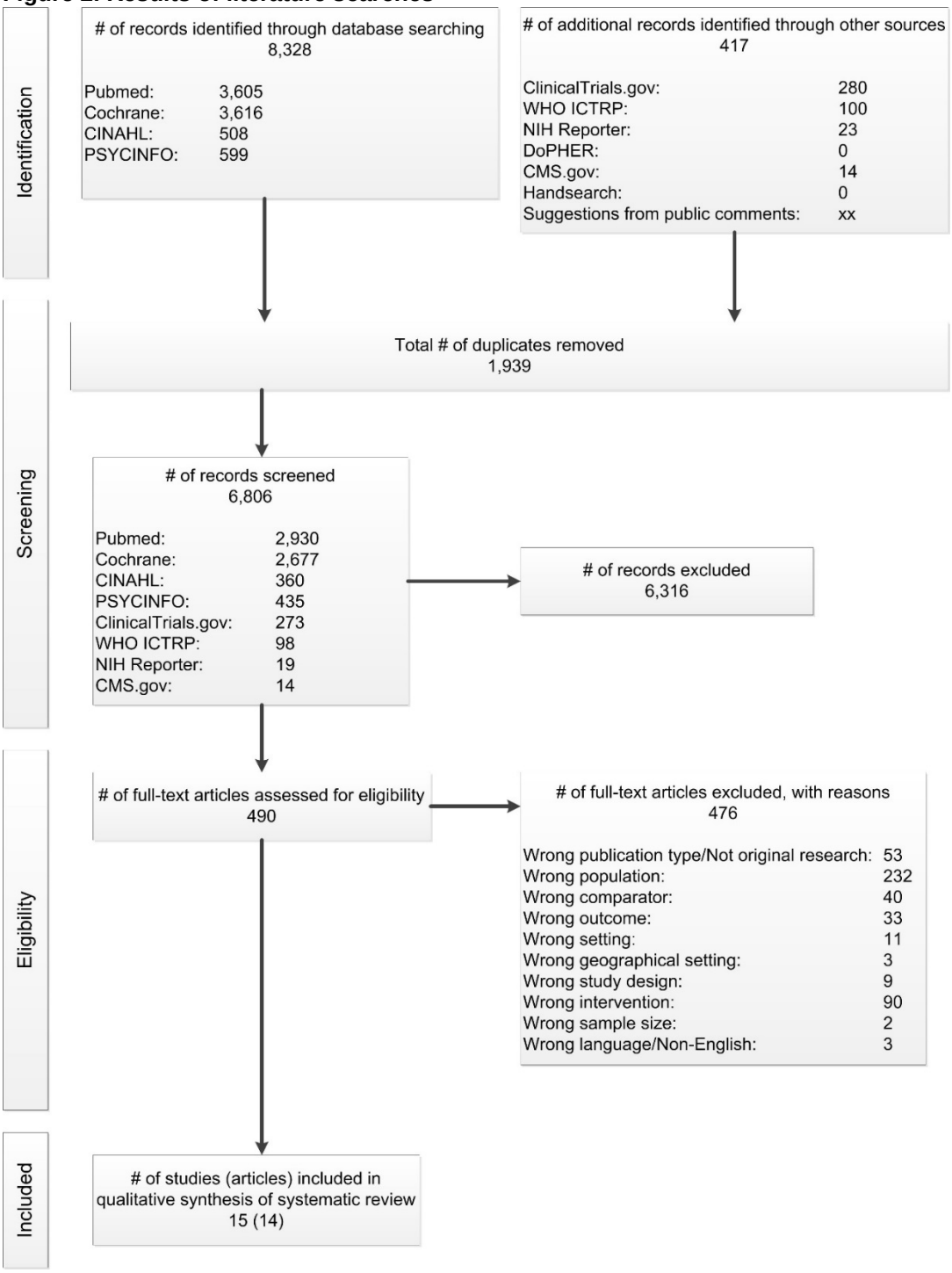
Introduction

This chapter presents the findings of this systematic review, starting with the results of the literature searches and description of included studies. The findings for each Key Question (KQ) present an overview of the identified evidence, followed by key points and detailed results. Detailed results include a description of relevant studies, intermediate outcome findings, patient health and service utilization outcomes, risk of bias considerations (with rating presented in full in Appendix C), and strength of evidence grades for each study. KQ 1 studies are presented individually. We synthesized the results qualitatively rather than quantitatively because of high levels of heterogeneity in the number and types of strategy components, differences between the experimental and control arms (i.e., in some studies, a single component distinguished strategy and control arms, and others, several components differed between arms), and outcomes assessed. We relied on author-reported measures of differences between groups and associated variances, but when these were not reported, we calculated differences and computed odds ratios (ORs) or mean differences, along with 95% confidence intervals (CIs) for between-arm comparisons.

Results of our searches appear in Figure 2. We reviewed 6,086 titles and abstracts dually and independently, and identified 490 articles for full-text review. Because of the lack of standard terminology used to define the types of studies of interest to this review, we used a wide-ranging search strategy. As a result, many citations were not relevant, leading to a much smaller pool of includes at full-text. We excluded 476 of these articles at full-text, leaving 14 articles representing 15 studies. Common reasons for exclusion included not meeting review criteria for population (i.e., not focusing on health care systems, organizations, or practitioners that provide mental health care for children and adolescents with mental health problems [n=232]), not meeting review criteria for comparator (i.e., not including a comparator [n=40]), not meeting review criteria for intervention (i.e., dissemination, implementation, or quality improvement (QI) strategies [n=90]), not meeting review criteria for publication type (cross-sectional studies, non-systematic reviews [n=53]), and not meeting review criteria for outcome (included only patient health outcomes or only intermediate outcomes for strategies not implementing an evidence-based practice [n=33]).

All full-text studies had a minimum of two independent reviewers, but for several studies, applying the inclusion/exclusion criteria consistently and reliably required multiple iterations of full-text review, often culminating in group discussions with the entire team to reach consensus. Our challenges arose from the process of applying consistent logic when encountering new or unanticipated "boundary" cases, the complexity of included studies, and the inadequacy of reporting in some instances. Several studies that were otherwise eligible were eventually excluded because they examined efficacy or effectiveness instead of the impact of a dissemination, implementation, or QI strategy.[56-60] In other instances, we excluded studies otherwise eligible for having the wrong comparator: we could not distinguish the effects of the strategy of interest (QI, implementation, or dissemination) from the underlying evidence-based practice (EBP).[61-67] We encountered strategies that used teachers and non-mental health care practitioners that we judged to be ineligible because they were not providers of mental health care.[68, 69] Some studies had relevant strategies but were directed at children who were *at risk of* but were not *identified with* mental health problems.[70] One study also required contacting the authors to obtain additional details on the care received by usual care participants, which was

Figure 2. Results of literature searches



unclear in the published article.[61-64] We did not identify any relevant studies upon our review of the English language abstracts of non-English language studies. A complete list of articles that were excluded during full-text review can be found in Appendix B.

Fourteen published articles reporting on 15 studies were identified that met the review inclusion criteria. One article included 2 studies within the same publication.[71] All 15 included studies (14 articles) addressed KQ 1, with six studies (5 articles) examining strategies classified

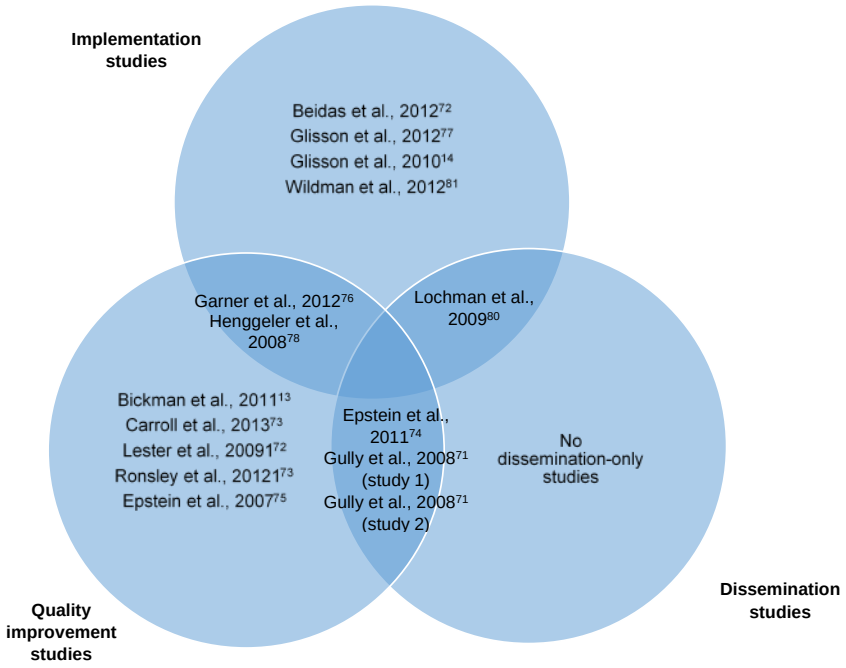
as professional training studies and 9 studies examining strategies classified as financial or organizational change studies based on categorizing components based on the EPOC taxonomy. One of the included articles addressed KQ 2, and 4 articles were identified that addressed KQ 3. The evidence base included randomized controlled trials,[13, 14, 71-79] controlled clinical trials,[80, 81] ecological aggregate,[82] and quasi-experimental.[71] Additional details describing the included studies are provided in the relevant sections of this results chapter and http://srdr.ahrq.gov/projects/530.

We assessed the risk of bias for all 15 studies dually and independently. We assessed 6 studies as high risk of bias. We assessed 2 studies as medium risk of bias. One study was rated as low risk of bias, and the risk of bias for 6 studies was determined to be unclear. Additional details describing the risk of bias assessment of all included studies is provided in Appendix C.

Description of Included Studies

Our review focused on 15 included studies from 14 publications.[13, 14, 71-82] Additionally, one of these studies addressed KQ 2 (harms) and four addressed KQ 3 (moderators of effectiveness). We first attempted to categorize each strategy by whether it focused on implementation, dissemination, or quality improvement according to our definitions, but we encountered several difficulties. First, our categorizations—assessed independently by two reviewers, with conflicts resolved by consensus—did not always match the study authors' categorization (dissemination, implementation or quality improvement). Second, the complexity of several of the strategies meant that we could not assign studies to mutually exclusive categories for implementation, dissemination, or quality improvement. We judged that 6 of 15 studies could be classified as having dual categories (Figure 3). Third, studies within the same category (dissemination, implementation, or quality improvement) did not have sufficient similarities in strategy components to enable meaningful synthesis of findings.

Figure 3. Venn diagram representing study team's original classification of included studies



As a result of these difficulties, we decided to categorize each of the strategies according to the EPOC taxonomy, indicating each of the professional, organizational, financial, or regulatory components that were present. No included studies contained regulatory components. Because many of the comparison groups also contained several of these components, we marked the components contained in each treatment and comparison group (i.e., study arm). This allowed us to fully describe the numerous components that were being combined and tested in each strategy, as well as enabled us to determine whether the study arms differed by a single or multiple components. Appendix E provides the full table of EPOC taxonomy components contained in each strategy by study arm; Table 3 presents a summary version of this table with descriptions of strategy components and differences by study arms. Strategies with one or more financial or organizational components were classified as "financial or organizational change" strategies, and strategies with only professional components were classified as "professional training" strategies. These categories guided our qualitative synthesis.

Table 3. Summary table of strategies tested

Author *Study Arms*	Target Condition and Ages of Youth	QI/I/D	Primary[a] Category-zation	Number and Types of EPOC Components Included in Strategy	Differences Across Study Arms
Beidas et al., 2012[76] *Augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (3 arms)*	Anxiety 8–17 years	I	Professional Training	2 Professional components (distribution of educational materials in 1 arm, educational meetings in two arms, plus one "other" component in all 3 arms—weekly consultation via virtual conferencing platform for 3 months after training)	Single difference across arms in the method of training program delivery (distribution of program's educational materials delivered via the computer in one group vs. workshop with behavioral role play and small group activities in another group vs. workshop with didactic instruction only in routine professional training group)
Bickman et al., 2011[13] Weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners	General mental health problem (children who receive home-based mental health treatment) Mean=15 years	QI	Financial or Organizational Change	5 Professional components (distribution of educational materials, educational meetings, patient-mediated interventions, audit and feedback, and one other—individual support by phone or email) 1 Organizational structural component (quality monitoring)	Single difference across arms (frequency of quality monitoring mechanism—weekly and cumulative 90-day vs. cumulative 90-day feedback to practitioners only)
Carroll et al., 2013[78] *Computer decision support plus EHR that included diagnosis and treatment guidelines vs. computer decision support plus EHR only*	General mental health problem (children who receive home-based mental health treatment) Mean=15 years	QI	Financial or Organizational Change	2 Professional components (patient-mediated intervention and reminders) 1 Organizational structural component (quality monitoring)	All components differed across arms (computer decision support plus EHR only group included none of these components)
Epstein et al., 2011[74] *Internet portal providing practitioner access to practice guidelines vs. wait-list control*	ADHD 6 to 12 years	QI/D	Financial or Organizational Change	5 Professional components (distribution of educational materials, educational meetings, patient-mediated interventions, audit and feedback, and reminders) 1 Financial component (provider incentives) 2 Organizational structural components (quality monitoring and staff organization)	All components differed across arms (wait-list control group includes none of these components)

Table 3. Summary table of strategies tested (continued)

Author *Study Arms*	Target Condition and Ages of Youth	QI/I/D[a]	Primary Category-zation	Number and Types of EPOC Components Included in Strategy	Differences Across Study Arms
Epstein et al., 2007[75] *Collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control*	ADHD Mean age=7 years	QI	Financial or Organizational Change	1 Professional component (audit and feedback) 1 Organizational provider oriented component (clinical multidisciplinary teams)	All (both) components differed across arms (control group included neither of these components)
Garner et al., 2012[72] *Paying practitioners for performance in successfully delivering of an EBP intervention vs. implementation as usual*	Substance use disorders; Mean age=16 years	QI/I	Financial or Organizational Change	4 Professional components (distribution of educational materials, educational meetings, educational outreach visits, one other—feedback from trained raters and weekly phone calls from developers of inter 2 Financial provider components (provider incentives and provider grant/allowance)	Single difference across arms in provider incentives
Glisson et al., 2012[73] *Organizational change vs. control*	General mental health problems 8–24 years	I	Financial or Organizational Change	5 Professional components (distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and one other—training and cognitive models to improve effectiveness) 1 Organizational provider-oriented component (satisfaction of providers with conditions of their work)	All components differed across arms (control group includes none of these components)

Table 3. Summary table of strategies tested (continued)

Author *Study Arms*	Target Condition and Ages of Youth	QI/I/D[a]	Primary Category-zation	Number and Types of EPOC Components Included in Strategy	Differences Across Study Arms
Glisson et al., 2010[14] *Organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (4 arms)*	Externalizing behaviors (youth referred to juvenile court with behavioral or psychiatric symptoms that require intervention) 6–17 years	I	Financial or Organizational Change	5 Professional components (distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and one other—training and cognitive models to improve effectiveness) 1 Organizational provider-oriented component (satisfaction of providers with conditions of their work) 1 Organizational structural component (quality monitoring)	All components differed across arms (control group included none of these components; MST groups were the only group with quality monitoring and audit and feedback; ARC groups were the only group with training and cognitive models to improve effectiveness and satisfaction of providers with conditions of their work)
Gully[b] et al., 2008[71] *Protocol to train nurses to educate parents about EBPs vs. typical services*	General mental health symptoms (children suspected of abuse during forensic medical examinations) 2–17 years	QI/D	Professional Training	4 Professional components (distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions)	All components differed across arms (typical services group includes none of these components)
Henggeler et al., 2008[80] *Intensive Quality Assurance system vs. workshop only to implement an EBP intervention*	Substance use disorders (adolescents with marijuana abuse) 12–17 years	QI/I	Financial or Organizational Change	4 Professional components (distribution of educational materials, educational meetings, educational outreach visits, one other—drug screen tests and supplies) 1 Financial provider component (other-money to facilitate treatment goals via a voucher system) 1 Financial patient component (patient incentives) 1 Organizational structural component (quality monitoring)	Two of 7 components differed across arms (patient incentives and quality monitoring were not part of the workshop only group)
Lester et al., 2009[79] *Professional training to identify and refer cases vs. usual care*	Psychosis (adolescents and adults with first-episode psychosis) 14–30 years	QI	Professional Training	4 Professional components (educational meetings, local consensus process, educational outreach visits, marketing)	All components differed across arms (control group included none of these components)

Table 3. Summary table of strategies tested (continued)

Author *Study Arms*	Target Condition and Ages of Youth	QI/I/D[a]	Primary Category-zation	Number and Types of EPOC Components Included in Strategy	Differences Across Study Arms
Lochman et al., 2009[77] *Professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (3 arms)*	Externalizing behaviors (children at risk for aggressive behaviors) Third-grade students	D/I	Professional Training	5 Professional components (educational materials, educational meetings, audit and feedback, marketing, one other—monthly ongoing training sessions)	Difference across the three study arms varied (training plus feedback group had all 5 components, training only group had educational meetings and marketing components, and control group had none of these components)
Ronsely et al., 2012[82] *Patient medication monitoring training program for practitioners vs. usual care*	Psychosis <19 years (mean=11)	QI	Professional Training	5 Professional components (educational materials, educational meetings, educational outreach visits, reminders, and one other—online access and project coordinator)	All components differed across arms (usual care group included none of these components)
Wildman et al., 2009[81] *Co-location of an EBP program in primary care vs. enhanced referral to an EBP program*	Externalizing behavior problems 2–12 years	I	Financial or Organizational Change	2 Organizational structural components (changes in scope and nature of benefits and services and one other—choice of treatment)	Single difference across arms (choice of treatment was not included in the enhanced referral group)

[a] original categorization made by study team

[b] applicable to both studies included in this publication

ADHD = Attention Deficit Hyperactivity Disorder; ARC = Availability, Responsiveness and Continuity; EBP = evidence-based practices; EPOC = Effective Practice and Organisation of Care; EHR = Electronic Health Record; MST = multisystemic therapy; QI/I/D = Quality Improvement\Implementation\Dissemination; vs. = versus.

Table 4 exhibits study characteristics of included studies organized by primary component of strategy according to the EPOC taxonomy (i.e., professional training or financial or organizational change).

Six studies had unclear risk of bias, 1 had low, 2 had medium, and 6 (in 5 publications) had high. Most studies were randomized controlled trials (RCTs) or controlled clinical trials. The majority were clustered at the practitioner, practice, or organizational level. The other two studies included a quasi-experimental and an ecological aggregate study. Six of the studies (in 5 articles) focused on professional training (i.e., only included professional components), while the other 9 studies focused on financial or organizational changes (i.e., included at least one financial or organizational component). Settings included primary care, community health, and schools. Each included study is reported in detail by KQ below. Full evidence tables are available at http://srdr.ahrq.gov/projects/530.

Table 4. Strategies to improve mental health of children and adolescents: Study characteristics

Study Descriptor	Characteristics	Primary Strategy: Professional Training[a]	Primary Strategy: Financial or Organizational Change[b]	Total
Design	RCT	2	0	2
	2-stage RCT	0	1	1
	Cluster RCT	2	6	8
	CCT	0	2	2
	Non-RCT	2	0	2
Setting	Primary Care	1	1	2
	Community Mental Health	4	8	12
	School	1	0	1
Primary Strategy Component	Professional Training[a]	6	0	6
	Financial or Organizational Change[b]	0	9	9
Strategy Categorization[c]	Quality Improvement	2	3	5
	Implementation	1	3	4
	Dissemination	0	0	0
	Hybrid QI and I	0	2	2
	Hybrid QI and D	2	1	3
	Hybrid I and D	1	0	1
Risk of Bias	Low	1	0	1
	Medium	0	2	2
	High	3	4	6
	Unclear	2	3	5
Key Question	KQ 1	6	9	15
	KQ 2	1	0	1
	KQ 3	1	3	4
Total N of studies		6	9	15

[a] Included all professional components

[b] Included at least 1 financial or organizational component

[c] Categories dually assigned by members of the study team according to the definitions of QU, I, and D included in the PICOTS

ADHD = attention deficit hyperactivity disorder; CCT = controlled clinical trial; D =Dissemination; I=Implementation; KQ = Key Question; QI = quality improvement; RCT = randomized controlled trial.

Key Question 1. Effectiveness of Strategies

Each of the 15 included studies from 14 publications[13, 14, 71-82] addressed our first KQ on strategy effectiveness. Key points and additional details of intermediate and patient health and service utilization outcomes overall and by the primary component of the strategy according to our classification based on the EPOC taxonomy (Table 5).

Table 5. Intermediate and patient health and service utilization outcomes by primary component of strategy

Outcome Category	Outcomes	Professional Training[a] (6 Studies in 5 Publications)	Financial or Organizational Change[b] (9 Studies/ Publications)	Total
Intermediate outcome: practitioner	Satisfaction/acceptability	1	0	1
	Adherence/fidelity	2	6	8
	Competence/skills	1	2	3
	Morale/engagement/stress	0	1	1
Intermediate outcome: system	Feasibility	0	0	0
	Uptake	0	0	0
	Timelines	0	0	0
	Penetration	0	0	0
	Sustainability	0	0	0
Intermediate outcome: patient	Resources (including costs)	0	0	0
	Access to care	3 in 2 publications	1	4 in 3 publications
	Satisfaction	2 in 1 publication	0	2 in 1 publication
	Treatment engagement	2 in 1 publication	0	2 in 1 publication
	Therapeutic alliance with provider	2 in 1 publication	0	2 in 1 publication
Patient health and service utilization outcome	Changes in mental health status	2	3	5
	Comorbid mental, substance use, developmental disorders	0	0	0
	Mortality	0	0	0
	Socialization skills and behaviors	1	0	1
	Functional status	0	1	1
	Quality of life	0	0	0
	Service utilization	2	1	3
Patient health and service utilization outcome not reported	N/A[c]	3 in 2 publications	4	7 in 6 publications
Total		6 in 5 publications	9	15 in 14 publications

[a] Included all professional components

[b] Included at least 1 financial or organizational component

[c] N/A= Not applicable because the strategy employed an EBP intervention, which has known benefits to these outcomes.

EBP = evidence-based practice.

Key Points: Characterization of Strategies

- We categorized the strategies tested in 6 studies (reported in 5 publications) as spanning multiple categories of our original three classifications: implementation, dissemination, or quality improvement. This overlap prompted us to use a different system, based on the EPOC taxonomy, to ultimately classify strategies as professional training (i.e., strategies that comprised only professional components) or financial or organizational change (i.e., strategies that comprised at least one financial or organizational component).
- We categorized 6 of the studies reported in 5 publications as examining professional training strategies and 9 of the studies as examining financial or organizational change strategies.

- The strategies tested all had multiple components, some of which spanned EPOC groupings of categories (e.g., professional, organizational, financial).
- The number of components differing between arms (treatment versus control) varied by study, with 4 studies having single component differences across arms (i.e., a single active component) and 11 studies with multiple component differences across arms (i.e., multiple active components).

Intermediate Outcomes

Practitioner Outcomes

We found studies that examined each of our three prespecified practitioner intermediate outcomes: satisfaction with or acceptability of approach (n=1), protocol adherence/program model fidelity (n=9), competence/skill (n=3)

- We found insufficient strength of evidence on practitioner satisfaction with or acceptability of approach from one study of professional training with three arms (augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP)
- For protocol adherence/program model fidelity intermediate outcomes, strength of evidence varied based on the specific strategy. We found
  - Low strength of evidence for benefit on three strategies. These included:
    - 1 professional training strategy (1 study, patient medication monitoring training program versus usual care) that had 5 active professional components (educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator as the active components)
    - 2 financial or organizational change strategies (1 study each)
    - 1 strategy on computer decision support plus electronic health record (EHR) diagnosis and treatment guidelines versus computer decision support plus EHR only that had 2 active professional components (patient-mediated intervention and reminders, and 1 organizational structural component, quality monitoring)
    - 1 Internet portal providing practitioner access to practice guidelines versus wait-list control that had 5 active professional components (distribution of educational materials, educational meetings, patient-mediated intervention, audit and feedback, reminders), 1 active financial component (provider incentives), and 2 active organizational structural components (quality monitoring and staff organization)
  - Low strength of evidence for no benefit on 2 strategies (1 study each). These included
    - 1 professional training strategy (augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP)
    - 1 financial or organizational change strategy (four arms: organization change and an EBP intervention versus an EBP intervention only versus organizational change only versus control)
  - Insufficient strength of evidence on 3 financial or organizational change strategies (1 study each). These included
    - 1 Intensive Quality Assurance system versus workshop only to implement an EBP intervention

    - 1 weekly and cumulative 90-dayversus cumulative 90-day feedback strategy to practitioners
    - 1 collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control
  - For practitioner morale/engagement/stress, we found low strength of evidence for benefit from one financial or organizational change study of organizational change versus control that had 5 professional components (distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and training and cognitive models to improve effectiveness) and 1 organizational provider-oriented component (satisfaction of providers with conditions of their work as the active components)
  - For practitioner competence/skill intermediate outcomes, strength of evidence varied based on the specific strategy. We found
    - Moderate strength of evidence for benefit from 1 financial or organizational change study of paying practitioners for performance of successful delivery of an EBP intervention versus implementation as usual that had a single active component, provider incentives
    - Low strength of evidence for no benefit from 1 practitioner change study of augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP
    - Insufficient strength of evidence from 1 financial or organizational change study of collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control

System Outcomes

Types of prespecified system intermediate outcomes included feasibility, uptake, timeliness, penetration, sustainability, and resources (including costs). No included study examined system-level intermediate outcomes.

Patient Outcomes

- Four types of prespecified patient intermediate outcomes were reported: patient access to care (n=4), satisfaction (n=2), treatment adherence (n=2), and therapeutic alliance (n=2).
- For patient access to care intermediate outcomes, strength of evidence varied based on the specific strategy. We found
  - Low strength of evidence for benefit on two strategies (3 studies reported in 2 publications). These included
    - 2 studies reported in one publication of a professional training strategy testing a protocol to train nurses to educate parents about EBPs versus typical services that had 4 professional components (distribution of educational materials, educational meetings, educational outreach visits, and patient-mediated interventions) as the active components
    - 1 study of a financial or organizational change strategy testing colocation of an EBP program in primary care versus enhanced referral to an EBP program that had a single active component, the organizational structural component of choice of treatment

    - Insufficient strength of evidence on a professional training strategy to identify and refer cases versus treatment as usual (1 study)
- We found low strength of evidence for benefit for each of the other three patient intermediate outcomes, patient satisfaction, treatment adherence, and therapeutic alliance. Two studies reported in 1 publication each examined a practitioner change strategy testing a protocol to train nurses to educate parents about EBPs versus typical services and contained 4 professional components (distribution of educational materials, educational meetings, educational outreach visits, and patient-mediated interventions) as the active components.

Key Points: Patient Health and Service Utilization Outcomes

- Six studies reported in 7 publications did not report a patient health or service utilization outcome because the strategy employed an EBP. For these studies, positive intermediate outcomes were assumed to have positive effects on patient outcomes. These included 2 professional training strategies tested in three studies, including
    - 1 professional training strategy testing augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP
    - 2 studies reported in one publication examining a protocol to train nurses to educate parents about EBPs versus typical services; and
    - 4 financial or organizational change strategies tested in 1 study each. These included:
        - 1 Internet portal strategy providing practitioner access to practice guidelines versus wait-list control
        - 1 colocation of an EBP program in primary care versus enhanced referral to an EBP program
        - 1 organizational change versus control
        - 1 Intensive Quality Assurance system versus workshop only to implement an EBP intervention).
- Five studies reported on changes in mental health status (symptoms, recovery, remission, etc.), one on socialization skills and behaviors, one on functional status, and three on service utilization. For mental health status outcomes, strength of evidence varied based on the specific strategy:
    - Low strength of evidence for benefit on a financial or organizational change strategy (1 study) of organizational change and an EBP intervention versus an EBP intervention only versus organizational change only versus control had for child behavior problem scores at 6 months (but had low strength of evidence for no benefit at 18 months) and low strength of evidence for benefit for out-of-home placements. The active components of the strategy tested in this study differed by arm (*Organizational change and an EBP intervention arm:* distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work, quality monitoring; *EBP intervention only arm:* distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, quality monitoring; o*rganizational change only arm:* distribution of educational materials, educational meetings, educational outreach visits, training and

cognitive models to improve effectiveness, satisfaction of providers with conditions of their work)
    - Low strength of evidence for no benefit on two strategies (from 1 study each) examining change in mental health status. These included
        - 1 professional training strategy of professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control
        - 1 financial or organizational change strategy testing paying practitioners for performance of successful delivery of an EBP intervention versus implementation as usual
    - Insufficient strength of evidence one two strategies (1 study each)
        - 1 professional training strategy to identify and refer cases versus treatment as usual
        - 1 financial or organizational change strategy testing collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control
- For changes in socialization skills and behaviors, we found low strength of evidence for no benefit of one strategy (from 1 study) testing professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control professional training strategy testing training plus feedback versus training only versus control.
- For changes in functional status outcomes, we found low strength of evidence for benefit of a financial or organizational change strategy (1 study) testing weekly and cumulative 90-day versus cumulative 90-day feedback to practitioners that had a single active component, a quality monitoring mechanism.
- The strength of evidence for service utilization outcomes varied based on the specific strategy. We found low strength of evidence for benefit on two strategies (1 study each). These studies examined changes in service utilization and included
    - 1 professional training strategy testing a patient medication monitoring training program versus usual care that had 5 active professional components (educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator)
    - 1 strategy to test professional training to identify and refer cases versus treatment as usual that had 4 active professional components (educational meetings, local consensus process, educational outreach visits, and marketing as the active components)
- In addition, we found insufficient strength of evidence for benefit of 1 financial or organizational change strategy (1 study), testing computer decision support plus EHR that included diagnosis and treatment guidelines versus computer decision support plus EHR only.

Detailed Synthesis

Professional Training Strategies

Six studies reported in five publications focused on strategies focused on professional training.[71, 76, 77, 79, 82] These studies each included various professional components according to the EPOC

taxonomy and no financial or organizational components. Studies included a maximum of five professional components. Components included distribution of educational materials, educational meetings, local consensus processes, educational outreach visits, patient-mediated interventions, audit and feedback, reminders, marketing, individual support by phone or email, project coordinator assistance, and monthly ongoing training sessions,

One strategy targeted school counselors, four targeted community-based mental health providers, and one targeted general practitioners. Two of these studies ultimately targeted children and adolescents with psychosis treated by general practitioners[79] or community-based mental health practitioners,[82] one targeted community therapists treating children with anxiety,[76] one targeted school counselors attempting to prevent externalizing behaviors among children at high risk of aggressive behaviors,[77] and two studies reported in one publication targeted children and adolescents who were suspected victims of abuse.[71] Details of each of these studies are described below.

Augmented Active Learning Versus Computerized Routine Versus Routine Professional Training Workshop to Implement an EBP Study Description

One RCT[76] (low risk of bias), conducted in 2009, focused on the implementation of CBT. Specifically, it evaluated the effectiveness of three 6-hour training modalities of CBT for anxiety in youth and the impact of ongoing consultation after training. Participants were 115 community therapists randomly assigned to one of three 1-day workshops to examine the effectiveness of training modality: routine training (RT, training as usual, n=41), computer training (CT, computerized training as usual, n=34), and augmented training that emphasized active learning (AT, n=39). After the workshops, all participants received three months of ongoing consultation that included case consultation, didactics, and problem-solving.

Participants included community therapists (mean age=35.9, 90 percent female) working in the community with children ages 8 to 17 with anxiety disorders who had not had more than 8 hours of prior CBT training for child anxiety. After training, all therapists received weekly consultation for 3 months. Outcomes measured at baseline and at posttraining included a measure of training satisfaction, a knowledge test of CBT for youth anxiety, and a measure of therapist fidelity that included adherence as assessed by six CBT competencies and skill of the therapist's competence as rated by a coder blind to treatment condition. Authors compared the means across treatment groups at posttreatment using t-tests; we calculated mean difference scores and 95% CIs of differences between the AT and RT groups and the CT and RT groups.

Therapists participated in an additional role-playing exercise at posttraining and 3-month followup (postconsultation) that involved simulated clinical situations where therapists encountered a research assistant acting as a child with anxiety seeking care. Independent assessors coded digital recordings of these sessions to determine the proportion of therapists in each training condition trained to adherence, skill, and knowledge criterion.

Table 6. Augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Beidas et al., 2012[76] RCT/Low	G1: Augmented active learning professional training workshop to implement an EBP intervention (CBT) G2: Computerized routine professional training workshop to implement an EBP intervention (CBT) G3: Routine professional training workshop to implement an EBP intervention (CBT) Single difference across arms: distribution of educational materials delivered via the computer in one arm vs. workshop with behavioral role play and small group activities in another arm vs. workshop with didactic instruction only in comparison arm.	G1: 40 G2: 34 G3: 41	Adherence to CBT: mean posttraining score difference G1–G3 and G2–G3 and differences in proportion of community therapists trained to criterion at posttraining (after a 1-day workshop) and postconsultation (after 3 months of followup consultation)	Mean post difference G1–G3: 0.43, 95% CI, -0.17 to 1.03 (calculated) Mean post difference G2–G3: -0.22, 95% CI, -0.89 to 0.45 (calculated) Posttraining G1 vs. G3 OR, 0.94, 95% CI, 0.39 to 2.30 G2 vs. G3 OR, 0.56, 95% CI, 0.21 to 1.48 Postconsultation G1 vs. G3 1.43, 95% CI, 0.55 to 3.73 G2 vs. G3 OR, 0.93, 95% CI: 0.35 to 2.49 (calculated)
			Skill/competence score: mean posttraining score difference between G1–G3 and G2–G3 and differences in proportion of community therapists trained to criterion at posttraining (after a 1-day workshop) and postconsultation (after 3 months of followup consultation)	Mean post difference G1–G3: -0.45, 95% CI, -1.10 to 0.20 (calculated) Mean post difference G2–G3: -0.46, 95% CI, -1.14 to 0.22 (calculated) Posttraining G1 vs. G3 OR, 2.00, 95% CI, 0.75 to 5.34 G2 vs. G3 OR, 0.68, 95% CI, 0.27 to 1.71 Postconsultation G1 vs. G3 OR, 0.94, 95% CI, 0.39 to 2.30 G2 vs. G3 OR, 0.97, 95% CI, 0.29 to 3.26 (calculated)
			Knowledge adherence to CBT: mean posttraining score difference between G1–G3 and G2–G3 and differences in proportion of community therapists trained to criterion at posttraining (after a 1-day workshop) and postconsultation (after 3 months of followup consultation)	Mean post difference G1–G3: -0.62, 95% CI, -1.45 to 0.21 (calculated) Mean post difference G2–G3: -0.03, 95% CI, -0.87 to 0.81 (calculated) Posttraining G1 vs. G3 OR, 1.50, 95% CI, 0.24 to 9.49 G2 vs. G3 OR, 2.61, 95% CI, 0.26 to 26.3 Postconsultation G1 vs. G3 OR, 1.07, 95% CI, 0.06 to 17.8 G2 vs. G3 OR, 0.39, 95% CI, 0.03 to 4.53 (calculated)

CBT = cognitive behavioral therapy; CI = confidence interval; EBP = evidence-based practice; G = group; N = number; OR = odds ratio; RCT = randomized controlled trial.

Intermediate Outcomes

All three modalities resulted in limited gains in therapist adherence, skill, or knowledge (Table 6). All groups improved in adherence to CBT measured by an Adherence Skills Checklist, participant skill (level of competence shown by the therapist in delivering treatment), and knowledge of CBT for youth anxiety, but there was no significant effect of training or interaction of time and training. In addition, the proportion of therapists trained to criterion did not differ across treatment groups for adherence, skill, or knowledge. Differences in satisfaction across training modalities were found ($F=7.22$, $df=2$ and 112, $p<0.001$), with participants in the CT group reporting lower satisfaction scores than the routine training group (50.8±5.9 versus

53.7±5.4; calculated mean difference, -2.9; 95% CI, -5.46 to -0.340; p=0.03. The study did not report statistically significant differences between the AT group and the RT group (55.5±4.8 versus 53.7±5.4; calculated mean difference, 1.8; 95% CI, -0.423 to 4.023; p=0.11).

Patient Health and Service Utilization Outcomes

The study did not report patient health and service utilization outcomes because the investigators implemented an EBT (CBT for anxiety).

Risk of Bias Considerations

This study was rated low risk of bias. Most procedures used by the authors in this study did not raise risk of bias concerns. One exception is that the authors randomized the therapist to treatment condition by date of enrollment, although allocation to group was concealed. Another minor concern was that the authors did not report differences in baseline characteristics between groups, and baseline differences, if significant, were uncontrolled in analyses. Attrition was 2 percent at posttraining and 12 percent at postconsultation (3-month followup) assessment.

Strength of Evidence

A single publication that presented data from an RCT yielded low strength of evidence that altering the modality of training community therapists to implement an EBT, CBT for youth anxiety, did not improve (1) practitioner satisfaction with or acceptability of approach, (2) protocol adherence or program model fidelity, or (3) skills or knowledge (Table 7).

Table 7. Augmented active learning versus computerized routine versus routine professional training workshop to implement an EBP: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Practitioner: satisfaction with or acceptability of approach for augmented training	1; 115 therapists	Low	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for no benefit calculated mean difference: 1.8, 95% CI, -0.423 to 4.023, p=0.11
Intermediate outcome: Practitioner: protocol adherence/program model fidelity	1; 115 therapists	Low	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for no benefit for various measures of uptake of training or scores of adherence
Intermediate outcome: Practitioner: competence/skills	1; 115 therapists	Low	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for no benefit for various measures of uptake of training or scores of adherence calculated mean difference: 1.8, 95% CI, -0.423 to 4.023, p=0.11

[a] Small sample size/number of events; CIs cross the line of no difference.

CI = confidence interval.

Protocol to Train Nurses to Educate Parents About EBPs Versus Typical Services Study Description

Two studies in one publication[71] examined a nurse-provided EBP intervention (high risk of bias) delivered to parents or caregivers to increase access to evidence-based mental health treatment for children ages 2 to 17 years who were suspected victims of abuse. The authors describe two studies; the first was quasi-experimental, and the second was an RCT. In Study 1, nurses at hospitals and community-based treatment centers received the EBP intervention, following which the investigators collected data from parents and caregivers. The comparison data came from the same study sites prior to the EBP intervention but from different parents and caregivers. In Study 2, parents or caregivers of children who were suspected victims of abuse presenting to a children's hospital forensic unit were randomly assigned to the EBP intervention protocol or typical services. The EBP intervention took place during forensic medical examinations performed for children who were referred for child abuse investigations. During the examination, nurses followed a protocol to educate parents and caregivers about EBPs for child mental health problems, addressing barriers to care and discussing with the parents the logistics of setting up an appointment.

In Study 1, the EBP intervention group included 172 parents or caregivers in both groups (number in each group not specified). In Study 2, the EBP intervention group included 24 parents or caregivers, and the usual care group included 27 parents or caregivers. Outcomes in both studies included parent/caregiver ratings (1–5, with 5 being *definitely* yes and 1 being *definitely no)* of access to EBT, satisfaction with services, helpfulness of mental health treatment, confidence to set and attend mental health treatment appointments, learning about an EBT, and feelings of being respected. Outcomes were assessed via phone calls with parents 1 month after the examination.

Intermediate Outcomes

Parents or caregivers in the EBP intervention group reported significantly higher ratings ($p<0.05$) than parents or caregivers in the usual care group for each of the six outcomes assessing access to EBTs, satisfaction with services, helpfulness of mental health treatment, confidence to set and attend mental health treatment appointments, learning about an EBT, and feelings of being respected. Outcomes were assessed via phone calls with parents 1 month after the examination (Table 8). The statistical significance of the findings held for both studies, with the exception that parents' or caregivers' ratings of the helpfulness of mental health treatment did not significantly differ between groups in Study 2 (i.e., the RCT). The authors conducted post-hoc analyses in Study 2 by examining data on ratings of access to EBTs categorically and only comparing those who answered definitely yes or definitely no across groups. The comparison maintained statistical significance ($X^2(1)=13.39$, $p=0.001$). Because randomization did not result in equal distributions of age and race/ethnicity across groups in Study 2, the authors also present hierarchical linear regression models adjusted for these covariates to compare ratings of outcomes between groups and found consistent patterns of significance across unadjusted and adjusted analyses.

Study authors also collected data on motivation to use the protocol and length of administration time from nurses in the EBP intervention group and combined across both studies. The nurses rated the protocol favorably and reported that the time to use the protocol took a mean of 2.4 minutes longer than it would have taken to address similar issues without the protocol.

Patient Health and Service Utilization Outcomes

The study did not report patient health and service utilization outcomes.[83, p. 8] The EBPs supporting the strategy included those identified as providing the "greatest level of theoretical, clinical, and empirical support" for abused children and their families: Trauma-Focused Cognitive Behavioral Therapy (TF-CBT), Abuse-Focused Cognitive Behavioral Therapy (AF-CBT), and Parent-Child Interaction Therapy (PCIT).[83]

Table 8. Protocol to train nurses to educate parents about EBPs versus typical services: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Gully et al., 2008[71] Study 1: Quasi-Experimental Design/High	Study 1: G1: Protocol to train nurses to educate parents about EBPs G2: Typical services All 4 professional components (distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions) differed across arms (comparison group strategy contained no components)	Study 1: G1: 86 G2: 86	Calculated mean G1–G2 difference in parent report (scale=1–5) of nurse discussing EBT during appointment	0.8 95% CI, 0.30 to 1.30
			Calculated mean G1–G2 difference in parent report (scale=1–5) of satisfaction with services	0.4, 95% CI, 0.15 to 0.65
			Calculated mean G1–G2 difference in parent report (scale=1–5) of perceived value of services	0.8, 95% CI, 0.52 to 1.08
			Calculated mean G1–G2 difference in parent report (scale=1–5) of confidence in setting/attending child mental health treatment appointments	0.9, 95% CI, 0.58 to 1.22
			Calculated mean G1–G2 difference in parent report (scale=1–5) of knowledge about EBPs	2.4, 95% CI, 2.04 to 2.76
			Calculated mean G1–G2 difference in parent report (scale=1–5) of rapport with the nurse	0.4, 95% CI, 0.15 to 0.65
Gully et al., 2008[71] Study 2: RCT/High	Study 2: G1: Education of parents and caregivers of children suspected of being abuse victims G2: Typical Services during a forensic medical examination All 4 professional components (distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions) differed across arms (comparison group strategy contained no components)	Study 2: G1: 24 G2: 27	Calculated mean G1–G2 difference in parent report (scale=1–5) of nurse discussing EBT during appointment	1.9, 95% CI, 1.13 to 2.67
			Calculated mean G1–G2 difference in parent report (scale=1–5) of satisfaction with services	0.9, 95% CI, 0.36 to 1.44
			Calculated mean G1–G2 difference in parent report (scale=1–5) of perceived value of services	0.6, 95% CI, -0.02 to 1.18
			Calculated mean G1–G2 difference in parent report (scale=1–5) of confidence in setting/attending child mental health treatment appointments	2.5, 95% CI, 1.86 to 3.14
			Calculated mean G1–G2 difference in parent report (scale=1–5) of knowledge about EBPs	1.1 95% CI, 0.60 to 1.60
			Calculated mean G1–G2 difference in parent report (scale=1–5) of rapport with the nurse	1.1, 95% CI, -0.49 to 1.69

CI = confidence interval; EBP = evidence-based practice; G = group; N = number; OR = odds ratio; RCT = randomized controlled trial.

Risk of Bias Considerations

Both studies presented in this publication were rated as having high risk of bias. Both studies had high attrition and no adjustment for missing data (48 percent for Study 1 and 41 percent for Study 2). In addition, Study 1 had the potential for confounding through nonrandom assignment, because of its quasi-experimental design. Study 2 did not report details about randomization, although the authors did conduct post-hoc tests and determined that neither variable that significantly differed between groups (age and race/ethnicity) was significantly associated with outcomes. All outcomes were parent or caregiver answers to nonvalidated questions measured on a Likert scale (1–5).

Strength of Evidence

A single publication presented data from two studies, a quasi-experimental study and an RCT. These studies yielded low strength of evidence that a nurse-led educational EBP intervention of parents of children suspected to be victims of abuse would increase patient access to care, satisfaction, treatment engagement, and therapeutic alliance with practitioner access to EBTs for mental health care, satisfaction, treatment engagement, and therapeutic alliance (Table 9).

Table 9. Protocol to train nurses to educate parents about EBPs versus typical services: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Patient: access to care	2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)	High	Consistent	Direct	Precise	Undetected	Low for benefit Strategy improved parent ratings of access to care (mean difference between groups ranged from 0.08 to 2.1 points in Study 1 and 0.6 to 1.9 in Study 2)
Intermediate outcome: Patient: satisfaction	2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)	High	Consistent	Direct	Precise	Undetected	Low for benefit Strategy improved parent ratings of satisfaction of care by a mean of 0.4 in Study 1 and 0.9 in Study 2
Intermediate outcome: Patient: treatment engagement	2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)	High	Consistent	Direct	Precise	Undetected	Low for benefit Strategy improved parent ratings of treatment engagement by a mean of 0.9 in Study 1 and 2.5 in Study 2
Intermediate outcome: Patient: therapeutic alliance with practitioner	2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)	High	Consistent	Direct	Precise	Undetected	Low for benefit Strategy improved parent ratings of therapeutic alliance by a mean of 0.4 in Study 1 and 0.9 in Study 2

RCT = randomized controlled trial.

Professional Training to Identify and Refer Cases Versus Usual Care

Study Description

One stratified cluster RCT[79] (high risk of bias) focused on testing an educational strategy targeting general practitioners (GPs). These GPs had access to early-intervention services for young people ages 14 to 30 with first-episode psychosis in three inner-city primary care trusts in Birmingham, England. Practices in the strategy arm received an educational intervention addressing practitioner knowledge, skills, and attitudes about first-episode psychosis; control practices did not receive the educational intervention but had access to the early-intervention services. The primary outcome was the difference in the proportion of young patients with first-episode psychosis referred to early-intervention services between practices. Secondary outcomes included duration of untreated psychosis, time to recovery (with recovery measured by the Positive and Negative Syndrome Scale PANSS), detainment under the Mental Health Act, and GP consultation rate. The authors used nonlinear mixed models to present the relative risk (RR) and 95% CIs for the primary outcome (difference in number of referrals per practice using Poisson error).

A total of 110 of 135 (81 percent) of eligible practices were recruited between 2004 and 2007 and randomized to the strategy or control arm (n=55 in each). One hundred seventy-nine patients with first-episode psychosis ages 14 to 30 were referred; 25 referred from the early-intervention services and 54 found eligible for inclusion during the team's audit of mental health notes. Eighty-three of the 179 patients provided secondary outcomes data (97 from the strategy group and 82 from the control group); a total of 68 of these were followed up 4 months later. Practices were recruited over three time periods, as more early-intervention services opened, allowing more practices to become eligible for study inclusion.

Intermediate Outcomes

The relative risk of referral to early-intervention services did not significantly differ between strategy and control practices (RR, 1.20; 95% CI, 0.74 to 1.95; p=0.48) (Table 10).

Patient Health and Service Utilization Outcomes

Several patient health and service utilization outcomes also did not differ between strategy practices and control practices (Table 10): detainment under the Mental Health Act during 4-month followup (14.9 percent versus 11.3 percent for intervention versus control practices, respectively, p=0.79), recovery at the end of 4-month followup as defined by a score of less than 10 on the positive subscale of the Positive and Negative Syndrome Scale (PANSS; 55.3 percent versus 64.4 percent, p=0.66 for strategyversus control practices, respectively), number of consultations in primary care (RR, 0.77; 95% CI, 0.45 to 1.33; p=0.34), and mean duration of untreated psychosis as retrospectively assessed at baseline (mean difference, -13.8; 95% CI, -199.1 to 171.6; p=0.88). Patients registered to strategy practices, however, had shorter delays than patients registered to control practices in reaching early-intervention services, as defined by the time from the first decision to seek care to the point of referral to an early-intervention service (mean difference, 222.03; 95% CI, 83.5 to 360.5; p=0.002).

Table 10. Professional training to identify and refer cases versus usual care: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Lester et al., 2009[79] Stratified cluster RCT/High	G1: Professional training to identify and refer cases G2: Usual care All professional components (educational meetings, local consensus process, educational outreach visits, marketing) differed across arms (comparison group strategy contained no components)	G1: 55 practices (97 patients) G2: 55 practices (82 patients)	RR for referral to early-intervention services after first contact	RR: 1.20 95% CI, 0.74 to 1.95 p=0.48
			Relative difference in detainment under the Mental Health Act within 4 months	Risk difference: 3.3%, p=0.79
			Duration of untreated psychosis as defined from onset of psychosis through receipt of early-intervention services	Mean difference = -13.8, 95% CI, -199.1 to 171.6, p=0.88
			Delay in reaching early-intervention services as defined from first decision to seek care through 4 months after strategy	Mean difference = -222.03; 95% CI, -83.5 to -360.5; p=0.002

CI = confidence interval; G = group; N = number; p = p-value; RCT = randomized controlled trial; RR = relative risk.

Risk of Bias Considerations

The study was rated high risk of bias due to high rates of patient attrition for the secondary outcomes (53.6 percent attrition for completion of study schedules and 62.0 percent attrition at 4-month followup). The authors did not use intention-to-treat models or adjust analyses for baseline differences across groups. For example, the randomization did not preclude the overrepresentation of young people from black and ethnic minority communities, but this difference was not accounted for in the analyses.

Strength of Evidence

A single publication that presented data from a stratified cluster RCT yielded insufficient evidence that a general professional training strategy to improve the identification and referral of cases of first-episode psychosis in young adults ages 14 to 30 changed early referral to care or mental health symptoms (Table 11). This study yielded low strength of evidence for benefit on duration of untreated psychosis.

Table 11. Professional training to identify and refer cases versus usual care: Strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Patient: access to care	1; 110 practices, 79 patients	High	Unknown (single study)	Direct	Imprecise[a]	Undetected	Insufficient RR of referral to early intervention: 1.20 95% CI, 0.74 to 1.95 p=0.48
Final outcome: change in mental health status	1; 158 patients for detainment under Mental Health Act, 83 patients for recovery	High	Unknown (single study)	Direct	Imprecise[a]	Undetected	Insufficient Patients in the professional training group did not have significant differences in change in mental health symptoms
Final outcome: Service utilization	1: 68 patients for number of consultations in primary care following the strategyand duration of untreated psychosis and delay in reaching early-intervention services	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit Patients in the professional training group averaged 223.8 days shorter for time from the first decision to seek care to the point of referral to an early-intervention service than patients in the control group

[a:] Small sample size/number of events; CIs cross the line of no difference

CI = confidence interval; RR = relative risk.

Professional Training Plus Feedback to Implement an EBP Intervention Versus Professional Training Only to Implement an EBP Intervention Versus Control (3 Arms)

Study Description

One cluster RCT study[77] focused on training school counselors to prevent the development of externalizing disorders among children at high risk for aggression. School counselors were trained to use the Coping Power (CP) program with third-grade children at high risk for aggressive behaviors as they transitioned to middle school. Counselors were randomly assigned to one of three study arms: CP training plus feedback (CP-TF), CP-basic training (CP-BT), or comparison; thus, the two groups testing strategies differed with respect to training intensity. School counselors in 57 public schools were randomly assigned to one of the three conditions, resulting in 19 schools per condition: 15 counselors in CP-TF, 17 in CP-BT, and 17 in the comparison group. Teachers nominated at-risk students using a rating scale for aggressive behavior. The upper 2 percent of aggressive students were excluded because they were believed

to be likely to have psychiatric diagnoses and were not appropriate for the indicated prevention programs. After screening 3,838 children, 1,422 met inclusion criteria. Contact was made with 670 of these potential participants, and 531 agreed to participate (79 percent). Sixty-five percent of the screened sample was boys. Eighty-four percent were African Americans, 14 percent were Caucasians, and 2 percent were of other race/ethnicity.

CP-TF provided more intensive training than CP-BT and had four components: (1) school counselors received three initial workshop training days in the fall, (2) school counselors participated in monthly ongoing 2-hour training sessions, (3) the trainers made individualized consultation to the school counselors available by email and telephone, and (4) the trainers reviewed the rate of completion of session objectives and provided individualized supervisory feedback through written and telephone contact to enhance EBP intervention integrity and received a monthly letter from their trainer. CP-BT had two training components that were the same as the first two components of CP-TF. School counselors received (1) 3 training days in the fall and (2) monthly 2-hour training sessions. Program delivery was assessed through eight variables evaluating concrete aspects of program delivery and counselor engagement in delivering the program. Seven measures were from audiotapes of child and parent sessions coded by research assistants using the objectives list for CP sessions after each EBP intervention session. Program delivery was assessed by measuring completion of EBP intervention objectives for the child and parent session measures of the number of sessions scheduled and number attended for children and parents. Ninety-four percent of the parents and children and 88 percent of teachers in the sample provided assessment data. The authors tested baseline differences between completers and those lost to followup within each of the three groups for each of seven outcomes and found significant differences on 3 of the 21 tests. In the comparison condition, those lost to followup had higher externalizing problems and lower social skills according to parents, but conversely they also had lower expectations that aggression would lead to better outcomes than completers in the comparison condition. Thus, the authors concluded that there were no clear patterns of differential attrition.

The authors used hierarchical linear modeling (HLM) to evaluate the effects of the strategy on post-strategy assessments of three externalizing behaviors and four positive targeted processes collected from children, caretakers, and/or teachers 2 years after baseline data collection. Intermediate outcomes included implementation outcome comparisons between CP-TF and CP-BT strategy group children. These included, for both children and for parents, number of sessions scheduled and attended, rates of strategy objectives completed, number of contacts with trainers, and ratings of counselors' engagement. Patient health and service utilization outcomes included children's externalizing behavior problems (as rated by teachers, parents, and children), positive social and academic behaviors as rated by parents and teachers, children's outcome expectations for aggressive behavior, and parents' consistency of discipline. Scores for participants in the training plus feedback (TF) group and basic training (BT) group were compared with scores of participants in the comparison group.

Intermediate Outcomes

The study did not examine differences in intermediate outcomes between groups. It did, however, present intermediate outcomes for differences between the two strategy groups that differed based on training intensity, which we considered to be relevant to KQ 3. We describe these comparisons, below, in the KQ 3 section.

Patient Health and Service Utilization Outcomes

The study found that children in the CP-TF group showed larger decreases in teacher-rated externalizing problems (p=0.01) and child-rated minor assaults (p=0.03), as well as larger increases in teacher-rated social and academic competence (p=0.01) than the control group children at 2-year followup (Table 12). No significant differences (p>0.05) on any of these outcomes were found between children in the CP-BT group and children in the comparison group.

Risk of Bias Considerations

This study was rated unclear risk of bias. The authors did not report details about the randomization method, allocation concealment, blinding of outcome assessors, or fidelity to protocol. Further, differences in baseline characteristics were not reported between groups so the success of the randomization method is unknown. In addition, analyses only adjusted for clustering within students, clustering between counselors, and for baseline levels of each outcome, so any significant differences in characteristics by group membership were uncontrolled. Although attrition was not at a concerning level (6 percent for parents and 12 percent for teachers), the authors noted some differences in characteristics between those lost to followup and completers in the methods section (data not reported). Furthermore, the authors did not use intention-to-treat analyses.

Strength of Evidence

A single publication that presented data from a cluster RCT focused on training school counselors to prevent the development of externalizing disorders among children at high risk for aggression yielded low strength of evidence for no benefit in improving either mental health symptoms or socialization skills and behaviors. This study examined the effect of altering the level of intensity in the dissemination of a prevention training and feedback program (Table 13).

Patient Medication Monitoring Training Program for Practitioners Versus Usual Care Study Description

One ecological aggregate study (outcomes were averaged across patients before and after the strategy)[82] evaluated the effect of a strategy to train practitioners to monitor patients on second-generation antipsychotics (SGAs) using a metabolic monitoring training program (MMTP). The MMTP instructed prescribers among the Vancouver Health Child and Youth Mental Teams on best practice in metabolic monitoring and the use of a metabolic monitoring and documentation tool (MMT). The study analyzed four data points before (from September 1, 2007, to December 31, 2008) and four after the implementation of the MMTP (from January 1, 2009, to April 20, 2010). The sample before MMTP implementation (N=1,114) was evaluated at baseline, 3, 6, and 12 months. The sample after MMTP implementation (N=1,262) was evaluated starting from a baseline measure (immediately after MMTP implementation) and then at 3, 6, and 12 months post-MMTP, using retrospective chart reviews.

Table 12. Professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Lochman et al., 2009,[77] cluster RCT/Unclear	G1: Professional training plus feedback to implement an EBP intervention (Coping Power:CP-TF) G2: Professional training only to implement an EBP intervention (Coping Power:CP-BT) G3: Control Training plus feedback arm had 5 professional components (educational materials, educational meetings, audit and feedback, marketing, one other—monthly ongoing training sessions), Training-only arm had 2 components (educational meetings and marketing), Control arm had no components	G1: 168 G2: 183 G3: 180	Behavior problems: beta coefficient, SE, and p value for externalizing composite (teacher rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within- student variation, and between- counselor variation	G1 vs. G3: -0.41, SE=0.16, p=0.01 G2 vs. G3: 0.10, SE=0.16, p=0.52
			Behavior problems: beta coefficient, SE, and p value for externalizing composite (parent rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within- student variation, and between- counselor variation	G1 vs. G3: -0.23, SE=0.12, p=0.05 G2 vs. G3: -0.13, SE=0.11, p=0.26
			Behavioral problems: beta coefficient, SE, and p value for minor assault (child rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within-student variation, and between-counselor variation	G1 vs. G3: -0.25, SE=0.12, p=0.03 G2 vs. G3: 0.04, SE=0.11, p=0.70
			Targeted processes: beta coefficient, SE, and p value for social/academic competence (teacher rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within-student variation, and between-counselor variation	G1 vs. G3: 0.35 SE=0.13, p=0.01 G2 vs. G3: 0.24, SE=0.13, p=0.06
			Targeted processes: beta coefficient, SE, and p value for social composite (parent rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within- student variation, and between- counselor variation	G1 vs. G3: 0.06, SE=0.12, p=0.65 G2 vs. G3: 0.15, SE=0.12, p=0.21
			Targeted processes: beta coefficient, SE, and p value for Outcome Expectations Questionnaire (child rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within-student variation, and between-counselor variation	G1 vs. G3: -0.24, SE=0.12, p=0.05 G2 vs. G3: 0.05, SE=0.12, p=0.67
			Targeted processes: beta coefficient, SE, and p value for inconsistent discipline (parent rated) score change at 2 yrs in HLM analyses adjusted for baseline score, within- student variation, and between- counselor variation	G1 vs. G3: 0.03, SE=0.11, p=0.80 G2 vs. G3: 0.04, SE=0.11, p=0.75

BT = basic training; CBT = cognitive behavioral therapy; CP-BT = Coping Power-Basic Training; CP-TF = Coping Power-Training plus Feedback; CI = confidence interval; EBP = evidence-based practice; G = group; HLM = hierarchical linear modeling; N = number; OR = odds ratio; RCT = randomized controlled trial; RR = relative risk; TF = training plus feedback; yr = year.

Table 13. Professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control:: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Final outcome: change in mental health status	1; 511 patients	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit CP-TF and CP-BT groups had fewer behavioral problems as rated by teachers (beta=-0.41, SE=0.16, p=0.01). There were no significant differences in teacher ratings of behavioral problem for the CP-BT versus comparison group or for any comparisons of behavioral problems as rated by parents.
Final outcome: change in socialization skills and behavior	1; 511 patients	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit CP-TF had fewer minor assaults as reported by the child (beta=-0.25, SE=0.12, p=0.03) and social/academic competence as reported by the teacher (beta=0.35, SE=0.13, p=0.01) as compared to comparison groups. These differences were not significant for the CP-BT versus comparison groups, nor were any significant differences found between groups on social skills as rated by parents.

CBT = cognitive behavioral therapy; CI = confidence interval; CP-BT = Coping Power-Basic Training; CP-TF = Coping Power-Training plus Feedback; OR = odds ratio; RCT = randomized controlled trial.

Intermediate Outcomes

A chart review revealed that the MMT was in the chart of 38.3 percent of recipients of SGAs postimplementation. The study also reported a drop in the prevalence of SGA prescribing between from 15.4 percent in the pre-MMTP period to 6.4 percent in the post-MMTP period ($p<0.001$) (Table 14).

Table 14. Patient medication monitoring training program for practitioners versus usual care antipsychotics: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Ronsley et al., 2012[82] Ecological aggregate-retrospective chart review/Unclear	G1: 1,262 post-MMTP patients G2: 1,114 pre-MMTP patients All professional components (educational materials, educational meetings, educational outreach visits, reminders, and one other—online access and project coordinator) differed across study arms (comparison group strategy contained no components)	G1: Patient medication monitoring training program (MMTP) for practitioners G2: Usual care	Proportion with SGA prescription Timing unclear	G1: 172/1,114 (15.4%) G2: 81/1,262 (6.4%) Calculated OR: 0.376; 95% CI: 0.284 to 0.496, p<0.001

CI = confidence interval; G = group; MMTP = Metabolic Monitoring Training Program; N = number; OR = odds ratio; SGA = second generation antidepressant.

Patient Health and Service Utilization Outcomes

For SGA-treated patients, the authors also reported monitoring rates pre- and postimplementation for anthropometric measures (height, weight, waist circumference, and blood pressure) and blood work parameters (fasting glucose, insulin, total cholesterol, triglycerides, high-density lipoprotein cholesterol or low-density lipoprotein cholesterol alanine aminotransferase or aspartate aminotransferase, prolactin). The authors reported that all measures were statistically significantly (at $p<0.01$) between pre-MMTP and post-MMTP measures at baseline, 3 and 6 months, but not at 12 months. Our calculated OR, however, based on overall reported Ns (Appendix D) suggests a decline in OR over time for some but not all measures. Nearly all measures continue to have statistically significant effects over time.

Risk of Bias Considerations

Many key details pertaining to study design and conduct were not reported. For example, we could not ascertain the proportion of patients that were retained in the sample before and after MMTP. A key concern for ecological studies relates to unmeasured concomitant strategies or secular changes. The rates of metabolic monitoring declined over time in both the pre-MMTP and the post-MMTP period. The authors did not comment on reasons for a decline in the pre-MMTP period, but in the absence of an active strategy during this phase, it is unclear whether concomitant external changes explain the pre-MMTP trend toward lower monitoring. Without further information on the reasons for the decline over time in monitoring before program implementation, the large differences between the pre-MMTP period and the post-MMTP period cannot be confidently attributed to the strategyalone. We thus rated this as having unclear risk of bias.

Strength of Evidence

One observational single study with high risk of study limitations yielded low strength of evidence that establishing a metabolic monitoring training program increases practitioner adherence and appropriate service utilization (Table 15).

Table 15. Patient medication monitoring training program for practitioners versus usual care: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence	1; 2,376 patients	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit 38.3% of patients had MMT in the charts after program implementation; drop in the prevalence of SGA prescribing between from 15.4% in the pre-MMTP period to 6.4% in the post-MMTP period (p<0.001)
Patient health and service utilization outcomes: service utilization	1; 253 patients (82 before, 171 after)	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit Increased metabolic monitoring over time (level of change varies by type of monitoring)

MMTP = Metabolic Monitoring Training Program; SGA = second generation antidepressant.

Financial or Organizational Change Strategies

Nine studies focused on changing systems of care.[13, 14, 72-75, 78, 80, 81] These studies each included at least one financial or organizational component according to the EPOC taxonomy. Three strategies included financial components and eight included organizational components (two strategies included at least one financial and at least one organizational component). Financial components included patient incentives, provider incentives, provider grants/allowances, and use of money to facilitate treatment goals in the form of a patient voucher. Organizational components involving structural change included quality monitoring, staff organization, changes in scope and nature of benefits and services, and enhanced referrals and choice of treatment. Organizational components directed at provider change included clinical multidisciplinary teams and improved satisfaction of providers with conditions of their work. The target of these strategies ranged from primary care clinicians or pediatricians (n=2) to community-based mental health therapists (n=3) or substance use providers (n=1) and practitioners in private, for profit behavioral health organizations providing home-based treatment (n=1), practices (n=1), and organizations (n=1). These studies ultimately targeted patients with ADHD (n=3), externalizing problems (n=2), substance use problems (n=2), and general mental health problems (n=2). We provide further details for each strategy below.

Internet Portal Providing Practitioner Access to Practice Guidelines Versus Wait-List Control

Study Description

One cluster RCT[74] examined the impact of an ADHD assessment and medication management program, facilitated through an Internet platform to pediatricians in primary care practices. The authors reported that the program was based on the evidence base for the American Academy of Pediatricians guideline recommendations. The trial included 511 children and 49 pediatricians spread across eight practices. Each matched practice pair (created according to the number of pediatricians and percentage of patients with Medicaid) was randomly allocated to either the strategy group (n=4 practices) or the control group (n=4 practices). The strategy group received the strategy immediately, while the control group received it after 6 months. Intervention practices participated in four 1-hour sessions of training on the new system, including didactic lectures and office flow modification workshops. Practices were then given access to an ADHD Internet portal that allowed parents, teachers, and pediatricians to input information (e.g., rating scales) about patients, after which information was scored, interpreted, and formatted in a report style that was helpful for assessing and treating patients with ADHD. Physicians evaluated their practice behaviors quarterly and addressed underperforming areas. Investigators conducted chart reviews at baseline and at 6-month followup for evidence of documentation of five guideline-specific measures. Investigators compared the percentage change in patients for whom each physician used each guideline behavior between baseline and followup between the strategy and control group using intent-to-treat analysis using t-tests. They did not account for clustering, given the small number of practices (n=8). They did not report differences in baseline levels of each outcome.

Intermediate Outcomes

Physicians in the strategy group had a higher mean change in the proportion of using parent ratings for assessment than physicians in the control group (p=0.03), the proportion using teacher ratings for assessment (p=0.04), and the proportion using *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) ADHD criteria during assessment (p=0.03) (Table 16). Pediatricians in the strategy group had greater decreases in the change in the proportion of using outside practitioners for ADHD diagnosis (p=0.0001) and in the proportion using teacher ratings to monitor treatment responses (p=0.0003). The difference in the proportion of pediatricians using parent ratings to monitor treatment responses between the strategy and control group did not reach significance (p=0.07).

Patient Health and Service Utilization Outcomes

The study did not report any final health or utilization outcomes. The strategy used an evidence-based practice (clinical practice guidelines set forth by the American Academy of Pediatrics).[84, 85]

Table 16. Internet portal providing practitioner access to practice guidelines versus wait-list control: Summary of results

Study Design/Risk of Bias	Study Arms: Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Epstein et al., 2007[74]/Cluster RCT/Unclear	G1: Internet portal providing practitioner access to practice guidelines G2: Wait-list control All components (5 professional components—distribution of educational materials, educational meetings, patient-mediated interventions, audit and feedback, and reminders, 1 financial component—provider incentives, 2 organizational provider-oriented components—clinical multidisciplinary teams) differed across study arms (comparison group strategy contained no components)	G1: 4 practices, 27 pediatricians, 501 patients G2: 4 practices, 22 pediatricians, 245 patients	Difference in mean baseline to 6-month followup change in proportion of practitioners using parent ADHD ratings during assessment	18.1, 95% CI, 2.05 to 34.2
			Difference in mean baseline to 6-month followup change in proportion of practitioners using teacher ADHD ratings during assessment	16.6, 95% CI, 1.61 to 31.6
			Difference in mean baseline to 6-month followup change in proportion of practitioners using DSM-IV criteria during assessment	29.4, 95% CI, 5.98 to 52.8
			Difference in mean baseline to 6-month followup change in proportion of practitioners using an outside practitioner for ADHD diagnosis	-50.0, 95% CI, -70.5 to -29.5
			Difference in mean baseline to 6-month followup change in proportion of practitioners using parent ratings of ADHD to monitor treatment responses	23.2, 95% CI, -1.78 to 48.2
			Difference in mean baseline to 6-month followup change in proportion of practitioners using teacher ratings to monitor treatment responses	32.4, 95% CI, 12.1 to 52.7

ADHD = attention deficit hyperactivity disorder; CI = confidence interval; DSM-IV = *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G = group; N = number.

Risk of Bias Considerations

We rated this study as having unclear risk of bias. The eight participating practices were matched according to number of pediatricians in the practice and the percentage of children on Medicaid; however, the study did not clarify whether other important differences existed in the practitioners or the patients between these practices. In addition, the study did not clarify whether patients were blinded to the strategy or whether outcome assessors were blinded to the outcome status of participants. The trial reported insufficient information to judge risk of bias on some criteria. For example, differences in baseline characteristics between groups are not reported.

Presumably, there would be few differences between groups in an RCT, but we could not assess the success of the randomization in this study.

Strength of Evidence

A single-cluster RCT, on providing access to ADHD guidelines to pediatricians via an Internet portal increases practitioner adherence and appropriate service utilization, yielded low strength of evidence for intermediate outcomes of practitioner protocol adherence and program model fidelity (Table 17).

Table 17. Internet portal providing practitioner access to practice guidelines versus wait-list control: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Practitioner—Protocol Adherence/Program Model Fidelity	1; 746 patients	Medium	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for benefit Strategy appeared to improve 4 of 5 examined outcomes that measured practitioner protocol adherence/program model fidelity outcomes (range mean change in outcome between groups range from 16.6 to -50) but estimates were very imprecise, with large confidence intervals.

[a] Wide CIs

Weekly and Cumulative 90-Day Feedback of Patient Symptoms and Functioning to Practitioners Versus Cumulative 90-Day Feedback Only of Patient Symptoms and Functioning to Practitioners

Study Description

One cluster RCT[13] evaluated the addition of weekly feedback of patient mental health symptoms and functioning as rated by youth, caregivers, and practitioners to therapists (the Contextualized Feedback System) in addition to standard 90-day feedback. The study hypothesized that this weekly feedback addition would improve mental health treatment effectiveness in a private mental health treatment organization. Clinicians in the control group received only the 90-day feedback. The trial initially randomized 24 sites to the implementation group and 25 sites to the comparison group. Over 40 percent of the sites (21 sites in total, 11 experimental, 10 control) dropped out of the study. The Symptoms and Functioning Severity Scale (SFSS) was scheduled to be completed every 2 weeks, but at least one measure in the battery was scheduled to be collected every week. The actual rate of data collection was lower than planned (mean records per youth: 11±9.2; mean number of weeks participated: 16.5±13.6), resulting in 1,341 SFSS scales from 340 youth, 1,291 SFSS scales from 144 clinicians, and 935 SFSS scales from 383 caregivers.

Intermediate Outcomes

Although the study did not specify an intermediate outcome, it reported results pertaining to practitioner adherence (Table 18). Specifically, the trial was intended to be a 2X2 factorial design with feedback as described above or the provision of three Web-based modules. The study reported that only 31 (the authors note that N=31 was one-third of the sample but do not specify the denominator) of clinicians accessed the Web-based module before their first client, indicating failure of implementation of the module. The authors reported no statistical differences in reasons for attrition between experimental and control groups but provided no additional details.

Table 18. Weekly and cumulative 90-day feedback of patient symptoms and functioning to practitioners versus cumulative 90-day feedback only of patient symptoms and functioning to practitioners: Summary of results

Study Design/Risk of Bias	Study Arms: Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Bickman et al., 2011[13] Cluster RCT/High	G1: Weekly and cumulative 90-day feedback of patient symptoms and functioning (CFS) to practitioners G2: Cumulative 90-day feedback only of patient symptoms and functioning (CFS) to practitioners Single difference in the presence of the organizational structural component, frequency of quality monitoring across arms	G1: 13 sites, 167 youths, 169 caregivers, 64 clinicians G2: 15 sites, 173 youths, 214 caregivers, 80 clinicians Total scales analyzed (breakdown by trial arm NR): youth, N=1,341; clinicians, N=1,291; caregivers, N=935	Youth Estimated coefficient of membership in feedback group at baseline Estimated coefficient of slope (time in weeks) Estimated coefficient of interaction of membership in feedback group and slope	 0.02, SE: 0.10, p>0.005 -0.001, SE: 0.002, p<0.0001 -0.01, SE: 0.002, p<0.001
			Clinicians Estimated coefficient of membership in feedback group at baseline Estimated coefficient of slope (time in weeks) Estimated coefficient of interaction of membership in feedback group and slope	 0.10, SE: 0.10, p>0.005 -0.005, SE: 0.001, p<0.0001 -0.01, SE: 0.002, p<0.0001
			Caregivers Estimated coefficient of membership in feedback group at baseline Estimated coefficient of slope (time in weeks) Estimated coefficient of interaction of membership in feedback group and slope	 0.01, SE: 0.13, p>0.005 -0.003, SE: 0.002, p>0.05 -0.01, SE: 0.003, p<.0001

CFS = Contextualized Feedback System; G = group; N = number; NR = not reported; RCT = randomized controlled trial; RoB = risk of bias; SE = standard error.

Patient Health and Service Utilization Outcomes

The authors conducted an HLM that nested repeated measures within participants, youths within clinicians, and clinicians within sites. They estimated three HLMs for each respondent

type. The models accounted for repeated measures at unequal durations within and across respondents using restricted maximum likelihood estimation and baseline differences by race through inclusion of race in the model. The study reported no difference at baseline in SFSS between groups (Table 18). Over time (the slope coefficient in the model), youths and clinicians reported significant improvement in SFSS scores (effect size=0.30 and 0.17, respectively, variance not reported), but caregivers did not. All three groups reported a higher rate of change in improvement in SFSS in the experimental group (effect sizes of 0.18, 0.24, and 0.27 for youths, clinicians, and caregivers, respectively) as calculated by study authors using the HLM-estimated coefficients measured at the average length of presence in the Contextualized Feedback Systems.

Risk of Bias Considerations

As noted above, 21 of 49 sites dropped out of the study after randomization, increasing the potential for high risk of bias. The authors noted that they conducted intention-to-treat analysis but did not specify whether their analysis pertained to all randomized sites or all patients within randomized sites. Although the authors noted no statistically significant differences for attrition, they did not provide details to judge these differences. Additionally, the publication did not provide key details on study design and conduct such as blinding of patients and outcome assessors, method of randomization, allocation concealment, fidelity to protocol, and timing of outcome measurement. Thus, we rated this study as having high risk of bias.

Strength of Evidence

The high risk of study limitations, coupled with lack of contextual details for poor practitioner adherence, results in insufficient evidence to judge the effect of the strategy on practitioner adherence (Table 19). The study provided low strength of evidence of benefit of the strategy on symptoms and functioning severity.

Table 19. Weekly and cumulative 90-day feedback of patient symptoms and functioning to practitioners versus cumulative 90-day feedback of patient symptoms and functioning to practitioners only: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence	1; N of practitioners unclear	High	Unknown (single study)	Direct	Unknown[a]	Undetected	Insufficient two-thirds did not view web modules
Patient health and service utilization outcomes: symptoms and functioning severity	1; 340 youth,144 clinicians, 383 caregivers	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit Membership in the weekly feedback group increased the rate of decline in symptoms and functioning severity scale by 0.01

[a:] Precision cannot be calculated without N of practitioners.

N = number.

Computer Decision Support Plus Electronic Health Record (EHR) That Included Diagnosis and Treatment Guidelines Versus Computer Decision Support Plus EHR Only

Study Description

One cluster RCT[78] evaluated the enhancement of a computer decision support system to improve the quality of ADHD diagnosis and treatment patterns, across 4 clinics and 84 patients (2 clinics and 42 patients per group). The strategy group used a previously studied intervention, Child Health Improvement through Computer Automation (CHICA), with ADHD guidelines embedded in the computer decision support system. The comparison group used CHICA without embedded ADHD guidelines. In addition to adherence to ADHD guidelines before and after the strategy, the study evaluated changes in prescriber behavior and skills, including changes in medication-prescribing patterns and mental health referral rates.

Intermediate Outcomes

The study reported several measures of improvements in practitioner adherence to the use of guideline-based care (Table 20). Children in the computer decision support arm were more likely to have been assessed using formal diagnostic tools than children in the control arm (OR, 8.0; 95% CI, 1.6 to 40.6). The authors also reported higher but not statistically significant differences in medication-prescribing patterns, reassessment of symptoms, or mental health referral rates. They noted that the study was underpowered to measure those outcomes. The authors reported a significant increase in the number of reported ADHD symptoms at the time of diagnosis in three out of four symptom domains ($p<0.01$ in the three domains).

Patient Health and Service Utilization Outcomes

Based on reported study data, we calculated wide and overlapping confidence intervals around improvement in visits to mental health specialists in the strategy arm when compared with the control arm (calculated OR, 2.195; 95% CI, 0.909 to 5.303; $p=0.081$, reported p-value in study=0.054) (Table 20).

Risk of Bias Considerations

Some details pertaining to study design and conduct were not reported, such as randomization approach and allocation concealment, baseline differences in clinics, blinding of outcome assessors, and fidelity to the protocol, leading to an unclear risk of bias.

Strength of Evidence

One cluster RCT with medium study limitations yielded low strength of evidence that computer decision support increases practitioner uptake of guidelines for diagnostic assessment and practitioner skills in measuring ADHD symptoms at diagnosis (Table 21). It also provided insufficient evidence to judge the effect on the strategy on measures of practitioner competence (specifically on reassessment of symptoms at 3 months, adjustment of medications, mental health referral) and on service utilization (visits to mental health specialist).

Table 20. Computer decision support plus electronic health record (EHR) that included diagnosis and treatment guidelines versus computer decision support plus EHR only: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Carroll et al., 2013[78] Cluster RCT/Unclear	G1: "Computer decision support plus EHR that included diagnosis and treatment guidelines G2: Computer decision support plus EHR without diagnosis and treatment guidelines. All components (2 professional components—patient-mediated intervention and reminders specific to ADHD and 1 organizational structural component—quality monitoring specific to ADHD) differed across study arms (comparison group strategy contained no components)	G1: 2 practices (42 patients) G2: 2 practices (42 patients)	Use of formal diagnostic assessment	Adjusted OR, 8.0; 95% CI, 1.6 to 40.6 p-value not reported Adjusted for age, gender, race/ethnicity, and insurance
			Number of core ADHD symptoms noted G1 vs. G2 (*exact N and SD of symptoms not reported by authors)*	Estimated mean difference: Inattentive symptom (home): -2.1, reported p<0.05 Inattentive (school): -0.9 reported p<0.05 Hyperactive (home): -2.2 reported p<0.05 Hyperactive (school): -1.2 reported p=0.075
			Documented medication adjustments	G1=45% G2=33% Calculated OR, 1.652; 95% CI, 0.683 to 3.998; p=0.266; reported p=0.45
			Reassessment of symptoms at followup visit	G1=50% G2=33% Calculated OR, 2.00; 95% CI, 0.829 to 4.838; p=0.123, reported p=0.36
			Mental health referral	G1=74% G2=55% Calculated OR, 2.323; 95% CI, 0.928 to 5.817; p=0.072, reported p=0.09
			Mental health visit	G1=67% G2=48% Calculated OR, 2.195; 95% CI, 0.909 to 5.303; p=0.081, reported p=0.054

ADHD = attention deficit hyperactivity disorder; CI = confidence interval; G = group; N = number; OR = odds ratio; p = p-value; RCT = randomized controlled trial; SD = standard deviation.

Table 21. Computer decision support plus electronic health record (EHR) that included diagnosis and treatment guidelines versus computer decision support plus EHR only: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence through uptake of guidelines for diagnostic assessment	1; 84 patients	Medium	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for benefit[b] Adjusted OR: OR, 8.0; 95% CI, 1.6 to 40.6
Intermediate outcome: practitioner adherence through reassessment of symptoms at 3 months, adjustment of medications, mental health referral	1; 84 patients	Medium	Unknown (single study)	Direct	Imprecise[b]	Undetected	Insufficient No statistically significant improvement on any measure
Intermediate outcome: practitioner adherence through measuring ADHD symptoms at diagnosis	1; 84 patients	Medium	Unknown (single study)	Direct	Imprecise[c]	Undetected	Low for benefit More reporting of 3 of 4 symptom domains
Final health outcome: service utilization (visit to mental health specialist)	1; 84 patients	Medium	Unknown (single study)	Direct	Imprecise[d]	Undetected	Insufficient Calculated OR: 2.195; 95% CI, 0.909 to 5.303; p=0.081, reported p-value in study=0.054

[a:] Small sample size/number of events; wide CIs.

[b:] Large magnitude of effect.

[c:] Small sample size/number of events.

[d:] Small sample size/number of events; CI cross the line of no difference.

ADHD = attention deficit hyperactivity disorder; CI = confidence interval; OR = odds ratio p = p-value.

Intensive Quality Assurance System Versus Workshop Only to Implement an EBP Intervention

Study Description

One controlled clinical trial,[80] with arms assigned at the therapist level, evaluated an evidence-based intervention for adolescent marijuana abuse in community settings. The study tested the use of Intensive Quality Assurance (IQA) to promote therapist implementation of contingency management (CM) techniques in a community mental health center setting. CM techniques provide concrete rewards for behaviors incompatible with substance use. The practitioners were multisystemic therapy (MST) clinicians; MST is a family- and community-based evidence-based treatment for adolescents presenting serious clinical problems that place them at imminent risk of out-of-home placement. Thirty practitioners in the strategy group received IQA. CM protocols were integrated into the existing MST quality assurance system, which included four manualized components including treatment, supervision, expert consultation, and organizational support, as well as ongoing training (e.g., quarterly booster training). The IQA group received weekly expert case consultation on CM cases and had the improvement of CM skills and competencies incorporated into existing clinician development plans and received quarterly booster training. The control group clinicians received usual care, which included participation in a CM workshop and access to the materials needed to implement CM and phone access to a CM expert for consultation. Both IQA and control group practitioners could access $150 for each patient to facilitate treatment goals; however, while those in the control group could use these funds to facilitate any aspect of MST treatment, including but not limited to CM interventions, practitioners in the IQA group were restricted to using these funds only for the CM voucher system that rewarded patients for clean drug screens.

The study conducted complex analyses of linear and quadratic trajectories of change over time in ratings of therapist adherence, using youth and caregiver ratings on the CM Therapist Adherence Measure. This measure included five items on a 4-point scale to measure CBT techniques and four items on a 3-point scale to measure monitoring techniques. These items were measured on monthly intervals from each family, leading to clustering within clinicians. The analysis used hierarchical linear modeling to account for this clustering. Because the study reported only gamma values from these analyses and did not provide sufficient additional information to provide context, our interpretation of their results is limited to the study's reported p-values.

Intermediate Outcomes

IQA was more effective than control at increasing practitioner implementation of CM cognitive behavioral techniques in the short term, based on youth and caregiver reports (Table 22). The study also reported (in text) that, based on youth reports, the effect of IQA was sustained. However, based on reported p-values and the discussion in the study, these increases did not appear to be not sustained at 6 months for both youth and caregiver reports. The study reported no difference by arm for practitioner implementation of CM monitoring techniques at 4 months and did not conduct further analyses at 6 months.

Patient Health and Service Utilization Outcomes

The study did not report any patient health and service utilization outcomes but cited an earlier study for evidence of effectiveness of CM for improving youth substance abuse outcomes.[86]

Table 22. Intensive quality assurance system versus workshop only to implement an EBP intervention: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Henggeler et al., 2008,[80] CCT/Unclear	G1: Intensive Quality Assurance system to implement an EBP intervention (CM) G2: Workshop only to implement an EBP intervention (CM) Two components (a financial patient component—patient incentives and an organizational structural component—quality monitoring) differed across study arms (comparison group strategy contained no components)	G1: 18 G2: 12	Gamma for youth report of therapist use of CM through 4 months	Gamma =0.78, SE=0.36, p=0.01
			Gamma for youth report of therapist use of CM T through 6 months	Gamma =0.12, SE=0.14, p=ns, details NR
			Gamma for youth report of clinician use of CM monitoring at 4 months	Gamma =0.03, SE=0.04, p=ns
			Gamma for caregiver report of clinician use of CM monitoring at 4 months	Gamma =0.09, SE=0.10, p=ns
			Gamma for caregiver report of clinician use of CM at 4 months	Gamma =0.79 , SE=0.30, p=0.01
			Gamma for caregiver report of clinician use of CM T at 6 months	Gamma=0.18, SE=0.09, p=0.04
			Gamma for caregiver report of clinician use of CM T at followup	Gamma=0.05, SE=0.54, p=ns

CCT = controlled clinical trial; CM = contingency management; G = group; IQA = Intensive Quality Assurance; N= number; ns = not significant; SE = standard error.

Risk of Bias Considerations

The study provided insufficient information to judge risk of bias on most criteria. Although the authors note randomization at the level of the supervisor (N=5) and their teams (N=8), they did not provide enough information to judge whether the study was fully randomized. For example, they offered no details on sequence generation or allocation concealment. The authors did not specify whether they randomized the supervisors and teams together or separately. They noted that they replaced therapists who left the program within a month, leaving unanswered the question of whether or how the new therapists were trained. Other unclear aspects of study design and conduct include validity of inclusion and exclusion criteria, blinding of outcome assessors, fidelity of the intervention, potentially concurrent interventions, attrition rate, and potential for crossover or contamination.

Strength of Evidence

One trial, with unclear risk of bias, provided insufficient evidence to judge the effect of an IQA approach compared with a workshop-only approach to improve therapist fidelity (Table 23). The study provided no information on patient health and service utilization outcomes.

Table 23. Intensive quality assurance system versus workshop only to implement an EBP intervention: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence to CBT and monitoring techniques	1; 30 practitioners, N of caregiver and patient reports and monthly data points NR	High	Unknown (single study)	Direct	Imprecise[a]	Undetected	Insufficient Study does not provide sufficient detail to judge magnitude of effect

[a:] Small sample size/number of events, CIs cannot be calculated.

CBT = cognitive behavioral therapy; N = number; NR = not reported.

Collaborative Consultation Treatment Service to Promote the Use of Titration Trials and Periodic Monitoring During Medication Management Versus Control

Study Description

One study[75] examined the use of a collaborative consultative model to improve the use of titration trials and medication monitoring during medication maintenance for children with ADHD. The study randomized by clusters, specifically, by practice. Twelve pediatric practice groups were randomized to collaborative consultation strategy (6 practices, 25 pediatricians recruited but data available on 16 pediatricians who saw 59 patients) and control groups (6 practices, 27 pediatricians recruited but data available on 22 pediatricians who saw 87 patients). Pediatricians in the collaborative consultative service group learned how to use titration trials (to determine optimal dosage) and rating scales (to monitor medication efficacy and side effects during medication maintenance). The authors report that the control group practices did not have access to consultative model services but did not otherwise describe care in the control group. Nine pediatricians in the strategy group and five in the control group did not enroll any children. Although the study authors initially conducted hierarchical linear modeling studies, the variance components associated with physician and pediatrician were negligible (variance NR) and were therefore dropped from all further analyses.

Intermediate Outcomes

Regarding uptake of the strategy, more pediatricians in the strategy group used titration trials than in the control group (interaction, β=-0.283; SE, 0.09; $p<0.01$) (Table 24). Both groups improved medication monitoring over time, but the study reported that monitoring did not increase more among pediatricians in the strategy group than in the control group (details not reported).

Our calculated OR for differences in the proportion of pediatricians citing specific obstacles preventing the implementation of evidence-based treatments found lower odds in the strategy arm for all outcomes, but these were statistically significant for the cited obstacles of lack of access to medications and lack of time for titration trials only (Appendix D).

Table 24. Collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Epstein et al., 2011[75] Cluster RCT/High	G1: Consultative collaborative treatment service to promote the use of titration trials and periodic monitoring during medication management G2: Control Two components (a professional component—audit and feedback and an organizational provider-oriented component—clinical multidisciplinary systems) differed across study arms (comparison group strategy contained no components)	G1 (pediatricians): 16 G2 (pediatricians): 22	Use of titration trials (interaction term from multivariate analysis)	Interaction of group effect and time β=-0.283, SE, 0.09; p<0.01
			Use of medication monitoring	Time β=0.200; p<0.01, interaction of group and time β NR, but strategy group noted as not having greater increase
	G1: Consultative collaborative treatment service to promote the use of titration trials and periodic monitoring during medication management G2: Control Two components (a professional component—audit and feedback and an organizational provider-oriented component—clinical multidisciplinary systems) differed across study arms (comparison group strategy contained no components)	G1 (patients): 59 G2 (patients): 87	Mean scores for combined parent and teacher ratings of ADHD symptoms NR F score for decrease in combined parent and teacher ratings of ADHD symptoms differing between groups at 12 months	$F_{1,144}$=0.05, p=0.83
			Mean scores for combined parent and teacher ratings of ADHD symptoms NR F score for decrease in combined parent and teacher ratings of ADHD symptoms, group x time interaction	$F_{2,144}$=.44, p=0.65

ADHD = Attention Deficit Hyperactivity Disorder; G = group; N = number; NR = not reported p = p-value.

Patient Health and Service Utilization Outcomes

ADHD scores for children, as rated by parents and teachers (combined in a single analysis), did not differ by group (F score for decrease in combined parent and teacher ratings of ADHD symptoms for group x time interaction: $F_{2, 144}$ =0.44, p=0.65).

Risk of Bias Considerations

The study experienced substantial practitioner and patient attrition. For practitioners, the differential rate of engagement was 17.5 percent; 5 of 27 pediatricians in the control group and 9

of 25 in the strategy group did not enroll any children in the study. The study did not offer explanations for differences in the rate of engagement. For patients, of 146 participants selected for followup, 45 had data from all 3 data points. The remaining 101 participants had at least 1 missing data point. The authors used a missing-at-random analysis because their analysis found no differences in DSM-IV–defined ADHD symptomatology at baseline between children with missing data and those who had complete data. Nonetheless, the risk of bias from low practitioner engagement and missing patient data put the study at high risk of bias.

Strength of Evidence

One cluster RCT with high study limitations provided insufficient evidence that the collaborative consultation model improved practitioner adherence (uptake), practitioner competency (cited obstacles to implementation of EBPs), and ADHD symptoms (Table 25).

Table 25. Collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence (uptake)	1; 38 practitioners	High	Unknown (single study)	Indirect	Imprecise[a]	Undetected	Insufficient Interaction for uptake of titration trials β=-0.283; SE, 0.09; p<0.01 Uptake of medication monitoring trials: p=NS, details NR
Intermediate outcome: practitioner competency (cited obstacles to implementation of EBPs)	1; 38 practitioners	High	Unknown (single study)	Indirect	Imprecise[a]	Undetected	Insufficient Lower odds with overlapping confidence intervals of citing obstacles in 6 of 8 measures (2 reach statistical significance)
Patient health and service utilization outcomes: ADHD symptoms	1; 144 patients	High	Unknown (single study)	Indirect	Imprecise[a]	Undetected	Insufficient F score for decrease in combined parent and teacher ratings of ADHD symptoms for group x time interaction: $F_{2, 144} = 0.44$, p=0.65

[a:] Small sample size/number of events

ADHD = Attention Deficit Hyperactivity Disorder; EBP = evidence-based practice; NR = not reported; p = p-value; SE = standard error.

Paying Practitioners for Performance in Successfully Delivering an EBP Intervention Versus Implementation as Usual Study Description

One study[72] studied a pay-for-performance strategy (medium risk of bias) to improve treatment implementation for adolescent substance use disorders. This cluster-randomized trial

evaluated the use of a pay-for-performance initiative among 986 adolescent patients treated by 120 therapists working in 29 different community-based substance use disorder treatment organizations. Organizations were randomized to an implementation-as-usual control condition (IAU) or to a pay for performance (P4P) experimental condition. Therapists across all organizations delivered the same evidence-based treatment using the Adolescent Community Reinforcement Approach (A-CRA), and each organization received standardized funding, training, and coaching from the treatment developers. Therapists in the pay-for-performance condition received $50 each month that they demonstrated competence in A-CRA treatment delivery and $200 for each patient who received a specified number of treatment procedures and sessions previously determined to be associated with improved patient outcomes. A therapist-level and patient-level propensity score were used to adjust for biases due to the cluster randomized design (i.e., therapists clustered within organizations and patients clustered within therapists). Adjusted intent-to-treat regression models were used to model two implementation outcomes: (1) number of therapists meeting A-CRA competence (using a Poisson distribution), (2) whether each patient met target A-CRA (using a Bernoulli distribution), and one QI effectiveness outcome: (3) patient-level remission status as defined in the Cannabis Youth Treatment Study at 6 month followup (Dennis et al., 2004).[87] The authors report event rate ratios (for the number of therapists' outcome) and ORs for the other two outcomes, as well as 95% CIs for differences between groups.

Intermediate Outcomes

Therapists assigned to the P4P condition had significantly higher likelihood of demonstrating A-CRA competence than the IAU therapists (24.0 percent P4P versus 8.9 percent, IAU; event rate ratio, 2.24; 95% CI, 1.12 to 4.48) (Table 26). Patients working with therapists assigned to the P4P condition were more likely to get target levels of A-CRA treatment procedures and sessions as compared with patients of therapists assigned to the IAU condition (17.3 percent P4P versus 2.5 percent IAU; OR, 5.19; 95% CI, 1.53 to 17.62).

Patient Health and Service Utilization Outcomes

The study found no significant difference in the likelihood of remission for adolescents working with therapists in the P4P conditions versus those working with therapists in the IAU condition (41.8 percent versus 50.8 percent; OR, 0.68; 95% CI, 0.35 to 1.33) over a 6-month timer period (Table 26).

Risk of Bias Considerations

The study was rated medium risk of bias because of high rates of patient attrition (20 percent for the intermediate outcome of percentage of patients getting target levels of A-CRA treatment procedures and sessions and 49 percent for the patient health outcome of patient remission status).[72] In addition, blinding of outcome assessors to the outcome status of participants was unclear.

Strength of Evidence

A single publication that presented data from an RCT examining a pay-for-performance strategy seeking to improve the implementation of an EBT to treat adolescents with substance use disorders yielded moderate strength of evidence for intermediate outcomes and low strength of evidence for no benefit of patient health and service utilization outcomes (remission) (Table 27).

Table 26. Paying practitioners for performance in successfully delivering an EBP intervention versus implementation as usual: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Garner et al., 2012[72] Cluster RCT/Medium	G1: Paying practitioners for performance (P4P) in successfully delivering an EBP intervention (A-CRA) G2: Implementation of an EBP intervention (A-CRA) as usual Single difference in financial provider component (provider incentives) across study arms (comparison group strategy contained no components)	G1:14 organizations, 49 therapists G2:15 organizations, 49 therapists	Practitioner-Competence/Skills as measured by therapist-level demonstration of A-CRA competence (count)	Event rate ratio=2.24, 95% CI, 1.12 to 4.48 OR, 5.19; 95% CI, 1.53 to 17.62
		G1: 14 organizations, 45 therapists, 429 patients G2:15 organizations, 40 therapists, 507 patients	Patient - Access to Care as measured by receipt of target A-CRA	OR, 5.19, 95% CI, 1.53 to 17.62
		G1: 14 organizations, 41 therapists, 254 patients G2:15 organizations, 40 therapists, 346 patients	Final Health/Patient-Centered Outcomes- Mental Health Symptoms, syndromes, or disorders as measured by patient remission at 6-month followup	OR, 0.68; 95% CI, 0.35 to 1.33

A-CRA= Adolescent Community Reinforcement Approach; CI = confidence interval; G = group; N = number; OR = odds ratio; RCT = randomized controlled trial; RoB = risk of bias.

Table 27 Paying practitioners for performance in successfully delivering of an EBP intervention versus implementation as usual: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Practitioner: competence/skills	1; 49 therapists and 936 patients	Medium	Unknown (single study)	Indirect	Precise	Undetected	Moderate for benefit Therapists in the P4P group were over twice as likely to demonstrate implementation competence as compared to IAU therapists (event rate ratio=2.24; 95% CI, 1.12-4.48); patients in the P4P condition were over five times as likely to meet target implementation standards (i.e., to receive specific numbers of treatment procedures and sessions) as IAU patients (OR, 5.19; 95% CI, 1.53 to 17.62) but confidence intervals were wide.
Final outcome: change in mental health status	1; 600	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit Patients in the P4P condition did not have significantly different rates of remission at end of treatment than patients in the IAU condition,

CI = confidence interval; IAU = implementation-as-usual control condition; OR = odds ratio; P4P = pay for performance.

Organizational Change Versus Control

Study Description

Two studies, conducted by the same group of authors, evaluated the Availability, Responsiveness and Continuity (ARC) program to improve organizational climate. The ultimate goal of ARC was to increase implementation of evidence-based practices.[14, 73] ARC involves multiple activities—targeted at stakeholders, management, and service practitioners—that are all specifically designed to improve implementation. Specifically, an ARC specialist worked with others to plan and endorse implementation activities, communicate a vision for implementation efforts, set performance standards, facilitate information sharing, identify services barriers, and plan to remove service barriers. The initial study, published in 2010, used a 2X2 RCT design of an EBP intervention called multisystemic therapy (MST) and ARC, resulting in four arms: MST+ARC, MST only, ARC only, and usual care across 14 rural Appalachian counties. The strategy lasted for 12 months and evaluated youth behavior problems, youth placement, and therapist adherence.[14] A follow-on study of 26 community mental health programs for youth

from a multisite mental health service system randomized the programs to ARC or usual care but sustained the strategy for 18 months to allow sufficient time to observe organizational change.[73]

Intermediate Outcomes

The original study reported no differences (but did not provide details) on any measure of therapist fidelity (therapist adherence as reported by caregiver, therapist rating of supervisor, audiocoded therapist adherence) (Table 28).[14]

Table 28. Organizational change and an EBP intervention versus an EBP intervention only versus organizational change only versus control: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Glisson et al., 2010[14] Two-stage RCT/Medium	G3: EBP intervention (MST) only G4: Organizational change (ARC) plus EBP intervention (MST) 5 professional components (distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and one other—training and cognitive models to improve effectiveness) 1 organizational provider-oriented component (satisfaction of providers with conditions of their work)	G3: NR by arm G4: NR by arm N overall: 243 (76% of total assigned to MST)	Caregiver rating of therapist adherence on the 28-item MST Therapist Adherence Measure—Revised (TAM-R)	No differences in caregiver-reported MST therapist adherence between ARC and non-ARC conditions, details NR
		G3: NR by arm G4: NR by arm N overall: 144 (46% of 316 assigned to MST)	Rating of therapist adherence based on audio coding of tapes from therapy sessions (TAM-R)	No differences in audio-coded ratings of therapist adherence between ARC and non-ARC conditions, details NR
		G3: NR by arm G4: NR by arm 257 SAMs provided by 91% of the therapists N overall: 257 (91% of therapists)	Therapist rating of supervisor adherence based on Supervisor Adherence Measure (SAM)	No differences in SAM ratings of supervisor adherence between ARC and non-ARC conditions, details NR

Table 28. Organizational change versus control: Summary of results (continued)

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
	G1: Control G2: Organizational change (ARC) only G3: EBP intervention (MST) only G4: Organizational change (ARC) and EBP intervention (MST) All components (5 professional components—distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and one other—training and cognitive models to improve effectiveness, 1 organizational provider-oriented component—satisfaction of providers with conditions of their work, and 1 organizational structural component-quality monitoring) differed across study arms (comparison group strategy contained no components)	G1:NR G1 and G2:291 G3 and G4:305 G4:NR	Probability of entering an out-of-home placement in the 18-month followup period (using a binomial sampling model based on the Bernoulli distribution and a logit link function), β; 95% CI, p	G1: -0.657, -1.030 to -0.280, p=0.003 G2: -0.587, -1.173 to -0.001, p=0.050 G3: -0.751, -1.355 to -0.147, p=0.019 G4: 0.365, 1.304 to -0.574, p=0.413
			Adjusted mean of Child Behavior Checklist Total Problem T Scores at 6 months, (based on hierarchical linear modeling) coefficient; 95% CI, p	Constant: 60.840; 95% CI, 63.684 to 57.996; p<0.001 G1: 0.715; 95%, CI 0.796 to 0.634; p<0.001 G2: -0.094; 95% CI, -4.461 to 4.273; p=0.964 G3: -0.042; 95% CI, -2.221 to 2.137; p=0.968 G4: -3.390; 95% CI, -6.597 to -0.138; p=0.040
			Child Behavior Checklist Scores at 18 months	G1: 57.30 G2: 56.75 G3: 55.30 G4: 55.85

Table 28. Organizational change versus control: Summary of results (continued)

Study Design/Risk of Bias	Study Arms: Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Glisson et al., 2012[73] Cluster RCT/Unclear	G1: Organizational change (ARC) G2: Control All components (5 professional components—distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, and one other—training and cognitive models to improve effectiveness and 1 organizational provider-oriented component—satisfaction of practitioners with conditions of their work) differed across study arms (comparison group strategy contained no components)	G1: 13 programs (N of clinicians by arm NR) G2: 13 programs (N of clinicians by arm NR) Total n of clinicians: 197	Morale coefficient; 95% CI, p	G1 vs. G2: 4.761; 95% CI, 2.239 to 7.283; p=0.001
			Job satisfaction coefficient; 95% CI, p	G1 vs. G2: 2.338; 95% CI, 0.929 to 3.747; p=0.003
			Organizational commitment coefficient; 95% CI, p	G1 vs. G2: 2.322; 95% CI, 1.110 to 3.534; p=0.001
			Stress coefficient; 95% CI, p	G1 vs. G2: -1.095; 95% CI, -6.305 to 4.115; p=0.667
			Emotional exhaustion coefficient; 95% CI, p	G1 vs. G2: -0.085; 95% CI, -2.024 to 1.854; p=0.929
			Role conflict coefficient; 95% CI, p	G1 vs. G2: -1.555; 95% CI, -2.999 to -0.111; p=0.036
			Role overload coefficient; 95% CI, p	G1 vs. G2: 0.566; 95% CI, -1.420 to 2.552; p=0.561
			Engagement coefficient; 95% CI, p	G1 vs. G2: 1.591; 95% CI, 0.217 to 2.965; p=0.025
			Personalization coefficient; 95% CI, p	G1 vs. G2: 1.275; 95% CI, 0.298 to 2.252; p=0.013
			Personal accomplishment coefficient; 95% CI, p	G1 vs. G2: 0.398; 95% CI, -0.274 to 1.070; p=0.233
			Functionality coefficient; 95% CI, p	G1 vs. G2: 2.845; 95% CI, 0.356 to 5.334; p=0.027
			Growth and advancement coefficient; 95% CI, p	G1 vs. G2: 1.370; 95% CI, 0.170 to 2.570; p=0.027
			Role clarity coefficient; 95% CI, p	G1 vs. G2: 0.784; 95% CI, -0.292 to 1.860; p=0.145
			Cooperation coefficient; 95% CI, p	G1 vs. G2: 0.585; 95% CI, -0.259 to 1.429; p=0.166
			Rigidity coefficient; 95% CI, p	G1 vs. G2: -2.689; 95% CI, -4.684 to -0.694; p=0.011
			Centralization coefficient; 95% CI, p	G1 vs. G2: -1.874; 95% CI, -2.923 to -0.825; p=0.001

Table 28. Organizational change versus control: Summary of results (continued)

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
			Formalization coefficient; 95% CI, p	G1 vs. G2: -0.992; 95% CI, -2.103 to 0.119; p=0.077
			Proficiency coefficient; 95% CI, p	G1 vs. G2: 1.154; 95% CI, -0.903 to 3.211; p=0.258
			Responsiveness coefficient; 95% CI, p	G1 vs. G2: 0.305; 95% CI, -0.717 to 1.327; p=0.543
			Competency coefficient; 95% CI, p	G1 vs. G2: 0.720; 95% CI, -0.542 to 1.982; p=0.250
			Resistance coefficient; 95% CI, p	G1 vs. G2: -0.523; 95% CI, -3.194 to 2.148; p=0.689
			Apathy coefficient; 95% CI, p	G1 vs. G2: -1.105; 95% CI, -2.077 to -0.133; p=0.028
			Suppression coefficient; 95% CI, p	G1 vs. G2: 0.078; 95% CI, -1.536 to 1.692; p=0.921

ARC = Availability, Responsiveness and Continuity; CI confidence interval; G = group; MST = multisystemic therapy; N = number; n = number; NR = not reported; p = p-value; RCT = randomized controlled trial; SAM = Supervisor Adherence Measure. TAM-R = Therapist Adherence Measure—Revised.

The followup study evaluated the Organizational Social Context measure for 126 clinicians across 26 programs, with a primary domain for morale (including subdomains of job satisfaction and commitment); three domains for organizational climate, namely engagement (personalization of engagement and personal accomplishment), functionality (growth and advancement, clarity of role, cooperation), and stress (emotional exhaustion, role conflict, role overload); and three domains for organizational culture, namely rigidity (centralization, formalization), proficiency (responsiveness, competency), and resistance (apathy, suppression).[73] Although all domains were in the expected direction (with the exception of suppression), the study reported statistically significant findings for morale (4.761; 95% CI, 2.239 to 7.283), engagement (1.591; 95% CI, 0.217 to 2.965), functionality (2.845; 95% CI, 0.356 to 5.334), and rigidity only (2.689; 95% CI 4.684 to 0.694).

Patient Health and Service Utilization Outcomes

The original study found that out-of-home placement was lower for youth in the MST-only or ARC-only conditions, when compared with usual care, but not MST plus ARC (Table 28).[14] The adjusted relative odds of a youth entering an out-of-home placement in a county that received the ARC intervention was 56 percent when compared with the odds of out-of-home placement in a county that did not participate in the ARC intervention. The adjusted relative odds of a youth entering an out-of-home placement who received MST treatment was 47 percent when compared with the odds of out-of-home placement for youth who did not receive MST. The study found that the combined arm reduced child behavior problems (measured by the Child Behavior Checklist Total Problem T Scores) at 6 months but not MST or ARC only. In a

piecewise regression, the MST plus ARC arm had a higher rate of decline than other arms in child behavior problem scores over the first 6 months. Between 6 and 18 months, however, the MST plus ARC had a lower rate of decline. At 18 months, all differences in child behavior problems flattened across conditions (MST plus ARC: 55.85, ARC: 55.30, MST: 56.75, and control: 57.30). The followup study did not evaluate final patient outcomes.

Risk of Bias Considerations

The original study[14] had a rate of attrition over 20 percent. It also did not offer information on fidelity and outcome assessor blinding and has the potential for recall bias for out-of-home placement. These issues place the study at medium risk of bias. The followup study replaced 2 of 26 sites that were found ineligible. They did not report key details such as differences in baseline characteristics and controls for the potential differences from replacement, so it was not possible to judge the effect of this alteration to the outcomes.[73] This placed the followup study at unclear risk of bias.

Strength of Evidence

Two RCTs with medium study limitations provided evidence on intermediate and patient health and service utilization outcomes for ARC; however, only one study provides evidence on each outcome category (Table 29). One study offers low strength of evidence that ARC does not improve practitioner adherence to MST. A second offers low strength of evidence that ARC improves some measures of practitioner morale, engagement, and stress. One study offers low strength of evidence that ARC reduces rate of out-of-home placement when compared with a control arm but no added benefit of a combined MST and ARC arm. One study offers low strength of evidence that the combined MST and ARC arm improves child behavior problems at 6 months but not at 18 months.

Colocation of an EBP Program in Primary Care Versus Enhanced Referral to an EBP Program

Study Description

A single controlled clinical trial[81] of high risk of bias reported evaluated the effect of colocating behavioral health care in primary care, specifically the Positive Parenting Program (Triple P), which has been shown to be effective in clinical trials. The study assigned four community-based, hospital-affiliated primary care pediatric practices in northeastern Ohio to colocated behavioral parent training, provided at the primary care office (the active arm) or enhanced referral to behavioral parent training delivered in settings outside the primary care office (the control arm). Additionally, the study considered the inclusion of a usual care arm, from seven community-based, hospital-affiliated primary care pediatric practices where patients could be referred to the behavioral parenting program routinely offered at the hospital. Only one family in the usual care arm sought a referral, so the study authors did not include the usual care in the final analyses. Parents in the colocation strategy arm were more likely to be younger and unemployed than parents in the control (enhanced referral) arm. Parental mean age was 31.8 (30.7 in the strategy arm and 34.8 in the control arm). Thirteen parents (59.1 percent) in the colocated arm and 1 parent (11.1 percent) in the enhanced-referral condition were unemployed.

Table 29. Organizational change versus control: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: practitioner adherence to MST (therapist and supervisor)	1; variable by analysis	Medium	Unknown (single study)	Direct	NR	Undetected	Low for no benefit Details NR, but does not demonstrate improvements in any measure of adherence by strategy group
Intermediate outcome: practitioner morale, engagement, stress	1; 197 practitioners in 26 programs	Medium	Unknown (single study)	Direct	Imprecise[a]	Undetected	Low for benefit Trends toward improvement in all domains, but nonoverlapping CI for only some domains
Patient health and service utilization outcomes: out-of-home placement	1; 615 youth	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for benefit Lower rate of out-of-home placement for MST or ARC, but not a statistically significant added benefit of MST plus ARC compared with control (34%)
Patient health and service utilization outcomes: child behavior problem scores at 6 months	1; 567 caregivers of youth	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for benefit Lower child behavior problem scores for the MST plus ARC arm, but not ARC or MST only compared with control
Patient health and service utilization outcomes: child behavior problem scores at 18 months	1; 567 caregivers of youth	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit No difference across arms

[a:] Small sample size/number of events; CIs cross the line of no difference for some domains.

ARC = Availability, Responsiveness and Continuity; CI = confidence interval; MST = multisystemic therapy; NR = not reported.

Outcomes of interest included rate of attendance of first Triple P appointments by parents, number of session attended, parental ratings of child externalizing behavior using the Eyberg Child Behavior Inventory, parent self-rated positive and negative affect as rated by the Positive and Negative Affect Schedule, and self-rated dysfunctional parenting as rated by the Parenting Scale.

Intermediate Outcomes

The study found that parents in the strategy arm were more likely to attend their first scheduled Triple P appointment (condition n=43/11,213; 0.38 percent of patient visits) than in the enhanced referral condition (n=12/9,704; 0.12 percent of patients' visits; OR, 3.10; 95% CI, 1.63 to 5.89) (Table 30). Parents in the colocated arm attended fewer Triple P sessions, on average, than the control arm, although the confidence interval around the mean difference is wide and the difference is not statistically significant (mean number of sessions, 3.07 versus

4.08; mean difference, -1.01; 95% CI, -2.60 to 0.58; p>0.2). The study did not control or adjust for baseline differences in study arms.

Table 30. Colocation of an EBP program in primary care versus enhanced referral to an EBP program: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Wildman et al., 2012 [81]/ CCT/High	G1: Colocation of an EBP program (PPP) in primary care G2: Enhanced referral to an EBP program (PPP) Single difference in one organizational structural component (enhanced referrals and choice of treatment) across study arms (comparison group strategy contained no components)	G1: 11,213 G2: 9,704	Proportion attending first scheduled triple-P appointment	G1: 43/11,213 G2: 12/9,704 OR, 3.10; 95% CI, 1.63 to 5.89
		G1: 43 G2: 12	Mean number of sessions attended (SD)	G1: 3.07 (2.42) G2: 4.08 (2.71) Calculated mean difference: -1.01; 95% CI, -2.60 to 0.58; p>0.2

CCT =controlled clinical trial; CI = confidence interval; EBP = evidence based practice; G = group; N = number; OR = odds ratio; PPP = Positive Parenting Program (Triple P); SD = standard deviation.

Patient Health and Service Utilization Outcomes

The study did not report patient health and service utilization outcomes by study arms but used an evidence-based intervention (Triple-P).

Risk of Bias Considerations

This study was rated high risk of bias. Study results were not adjusted for baseline differences (colocation parents were older and unemployed compared with the control group parents) and between-practice differences in culture were not considered. Little information was given regarding how the practitioners explained the program to patients and whether there were other patient, primary care practitioner, or practice attributes that differed between groups. Whether all clinics were randomized to a condition, whether outcome assessors were blinded to condition, and levels of attrition also were unclear.

Strength of Evidence

A single publication that presented data from a controlled clinical trial of colocation of behavioral health care in primary care settings yielded low strength of evidence that altering the modality of training community therapists to implement an EBT (CBT for youth anxiety) increases access to care (Table 31).

Table 31. Co-location of an EBP program in primary care versus enhanced referral to an EBP program: Detailed strength of evidence

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Patient access to care (attending first Triple P visit)	1; 4 pediatric practices, 20,917 children with primary care visits	High	Unknown (single study)	Direct	Precise	Unclear	Low for benefit OR, 3.10; 95% CI, 1.63 to 5.89

CI = confidence interval; OR = odds ratio.

KQ 2 Harms

A single study, focused on general professional training to identify and refer first-episode cases of psychosis, reported harms.[79]

Key Points

- Only one study reported on the harms of strategies to improve mental health in children and adolescents.
- The study did not find any harms associated with a strategy to train general practitioners to improve identification of first-episode cases of psychosis.

Detailed Study Description

A single study that examined harms associated with a strategy to improve the mental health care of children and adolescents met our inclusion/exclusion criteria. The study,[79] described previously, was classified as a professional training strategy. This study found no harms associated with a practitioner education strategy to improve the rates of referral to early intervention services for first-episode psychosis experienced by young people ages 14 to 30 (Table 32). The investigators found no differences between strategy and control practices with respect to patients who reported adverse events (n=0 in both groups, details of specific adverse events measured not reported) and no increase in false-positive referrals from primary care before and during the study. The authors report that the rate of false-positive referrals within practices "remained between 12.7 percent and 13.4 percent before and during the study"; however, differences between groups were not reported.

Table 32. Harms associated with professional training to identify and refer cases versus usual care: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Lester et al., 2009[79] Stratified Cluster RCT/High	G1: Professional training to identify and refer first-episode cases of psychosis G2: Usual care All professional components (educational meetings, local consensus process, educational outreach visits, marketing) differed across arms (comparison group strategy contained no components)	G1: 55 practices (97 patients) G2: 55 practices (82 patients)	Adverse events within 4 months	0 reported in both groups
			False positive referrals from primary care	NR other than that there was no increase, remaining between 12.7% and 13.4% before and during the study.

CI = confidence interval; G = group; NR= not reported; RCT = randomized controlled trial

Strength of Evidence

A single publication with high risk of bias presented data from a practitioner education yielded insufficient evidence for patient harms (i.e., side effects including adverse events and false positive referral rates) (Table 33).

Table 33. Harms associated with professional training **to identify and refer cases versus usual care: Detailed strength of evidence**

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Intermediate outcome: Patient side effects	1; 110 practices, 179 patients	High	Unknown (single study)	Direct	Unknown[a]	Undetected	Insufficient No adverse events were reported, no significant between-group differences for false-positive referral rates from primary care

[a:] Insufficient data to calculate precision

KQ 3: Moderators

Four studies examined moderators of the effectiveness of strategies on outcomes. Three of these studies examined treatment intensity as a moderator of the effectiveness of professional training of school counselors to prevent children at high risk for aggressive behaviors from developing externalizing problems,[77] a financial or organizational change of adding weekly

feedback to therapists providing home-based mental health care for children and adolescents,[13] and a financial or organizational change of a collaborative consultative model to improve the use of titration trials and medication monitoring during medication maintenance for children with ADHD.[75] The other study examined whether fidelity to protocol improved the effectiveness of a pay-for-performance strategy to improve treatment implementation of an EBP for adolescent substance use disorders.[72]

Key Points

- The strength of evidence of the three studies that examined treatment intensity as a moderator varied from having grades of low for benefit (for the association between prevention of externalizing behaviors on improving patient mental health symptoms) to insufficient (for the other two studies that examined treatment intensity as a moderator of the effectiveness of a weekly feedback strategy to improve home-based mental health care and of a collaborative consultative strategy to improve the use of EBPs) due to myriad study limitations and inability to determine the precision of findings.
- The strength of evidence was graded as low for no benefit for the moderating effect of fidelity to the protocol on effectiveness of a pay-for-performance strategy to improve treatment implementation of an EBP for adolescent substance use disorders.

Detailed Study Description

Moderating Effects of Intervention Effects—Intensity of the Intervention

Study Description

Three studies address the moderating effects of higher intensity of the strategy on outcomes. One study[77] of unclear risk of bias examined the effectiveness of professional training of school counselors to use the CP program with third-grade children at high risk for aggressive behaviors as they transitioned to middle school. Counselors were randomly assigned to one of three conditions: CP-TF, CP-BT, or comparison; thus, the two strategy groups differed with respect to training intensity. CP-TF was more intense and included four components, while CP-BT included two training components. The findings indicated that the CP-TF group had greater decreases in externalizing behavioral problems as rated by teachers than the comparison group (mean difference= G1 versus G3: OR, 0.66; 95% CI, 0.41 to 0.91), greater decreases in child-rated minor assault than the comparison group (G1 versus G3: OR, 0.78; 95% CI, 0.62 to 0.99), and greater improvements in teacher-rated social/academic competence than the comparison group (G1 versus G3: OR, 1.42; 95% CI, 1.10 to 1.83). None of the outcomes significantly differed for the CP-BT and comparison groups. The authors then compared the outcomes for the CP-TF and CP-BT groups to determine whether improvements in outcomes differed by training intensity. The authors also compared the two strategy groups on provider protocol adherence/program fidelity and patient access to care and treatment engagement intermediate outcomes.

A second study[13] categorized as targeting a financial or organizational change evaluated the addition of weekly feedback to therapists (the Contextualized Feedback System) in addition to standard 90-day feedback on symptoms and functional status change of children receiving home-

based mental health treatment. Regarding intermediate outcomes, the investigators reported implementation failure in one arm. Regarding patient health and service utilization outcomes, the study reported a higher rate of change in improvement in SFSS in the experimental group (effect sizes of 0.18, 0.24, and 0.27 for youths, clinicians, and caregivers, respectively; calculated by study authors using the HLM-estimated coefficients measured at the average length of stay in the Contextualized Feedback Systems. The authors sought to understand the dose-response effect of the strategy, specifically, whether there was an association between the proportion of reports viewed and outcomes (symptoms and functional status).

A third study,[75] also categorized as targeting a financial or organizational change, examined the use of a collaborative consultative model to improve the use of titration trials and medication monitoring during medication maintenance for children with ADHD. The study found a higher rate of practitioner uptake of titration trials in the strategy arm and no effect (or no consistent effect) of the strategy on uptake of medication monitoring during the maintenance phase of the drug, practitioner competency (measured by cited obstacles to implementing EBPs), or ADHD symptoms. The study then sought to understand the effect of undertaking a titration trial on ADHD symptoms.

Intermediate Outcomes

The study on varying training intensity for school counselors[77] compared the CP-TF versus the CP-BT groups on practitioner protocol adherence and program model fidelity as well as patient access to care and treatment engagement intermediate outcomes (Table 34). The author report no significant differences in rates of child sessions scheduled; however, calculated mean differences indicated that the CP-TF group had fewer sessions scheduled for children and for parents than the CP-BT (calculated mean difference and 95% CI, -3.10; -3.60 to -2.60 for children and -0.50; -0.24 to -0.77 for parents). There were no significant differences between groups with respect to child and parent attendance, as well as parent treatment engagement. For children, however, treatment engagement was significantly better for the CP-TF group than for the CP-BT group (calculated mean difference and 95% CI, 0.30; 0.28 to 0.32). Practitioners in the CP-TF group also had a greater number of contacts with trainers than those in the CP-BT group (calculated mean difference and 95% CI, 18.10; 17.51 to 18.69).

Patient Health and Service Utilization Outcomes

The study on training intensity for school counselors demonstrated that the strategy effects were significantly different for the CP-TF and CP-BT HLM contrasts for teacher-rated externalizing behaviors over time (mean change for CP-TF=0, mean change for CP-BT=5, $X^2$2(1)=3.87, p=0.05), child self-reported assaultive behaviors (mean change for CP-TF=0.18, mean change for CP-BT=0.45, X^2 (1)=6.23, p=0.01), and child-rated expectations of the utility of aggression (mean change for CP-TF=-0.1, mean change for CP-BT=0.1, X^2 (1)=5.64, p=0.02) (Table 35).[77]

Table 34. Intensity of the strategy as a moderator of the effectiveness of the strategy: Summary of results (intermediate outcomes)

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Lochman et al., 2009[77]	G1: Professional training plus feedback to implement an EBP intervention (CP-TF) G2: Professional training only to implement an EBP intervention (CP-BT) G3: Control Difference across the three study arms varied: 2 strategy arms and 1 control arm—training plus feedback arm had all 5 components, training only arm had educational meetings and marketing components, and control arm had none of these components)	G1: 168 G2: 183 G3: 180	Calculated mean difference (and 95% CI) in rates of child sessions and parent sessions scheduled (G1-G2)	-3.10 (-3.60 to -2.60) child -0.50 (-0.24 to --0.77) parent
			Calculated mean difference (and 95% CI) in rates of attendance for child and parent sessions (G1-G2)	0.01 (-0.08 to 0.11) child -0.04 (-0.12 to 0.05) parent
			Calculated mean difference (and 95% CI) in number of strategy objectives completed for child and parent sessions (G1-G2)	0 (-0.21 to 0.21) child 0 (-0.21 to 0.21) parent
			Calculated mean difference (and 95% CI) in number of contacts of practitioners with trainers (G1-G2)	18.1 (17.51 to 18.69) practitioner
			Calculated mean difference (and 95% CI) in ratings of counselors' engagement with children and with parents (G1-G2)	0.30 (0.28 to 0.32) children -0.10 (-0.12 to -0.08) parent

CI = confidence interval; CP-BT = Coping Power-Basic Training; CP-TF = Coping Power-Training plus Feedback; G = group; N = number.

Table 35. Intensity of the strategy as a moderator of the effectiveness of the strategy: Summary of results (patient health and service utilization outcomes)

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Lochman et al., 2009[77]	G1: Professional training plus feedback to implement an EBP intervention (CP-TF) G2: Professional training only to implement an EBP intervention (CP-BT) G3: Control Difference across the three study arms varied: 2 strategy arms and 1 control arm—training plus feedback arm had all 5 components, training only arm had educational meetings and marketing components, and control arm had none of these components)	G1: 168 G2: 183 G3: 180	Mean change in teacher-rated externalizing behaviors over time	CP-TF=0, CP-BT=5, $X^2$2(1)=3.87, p=0.05
			Mean change in child self-reported assaultive behaviors	CP-TF=0.18, CP-BT=0.45, X^2 (1)=6.23, p=0.01
			Mean change for child-rated expectations of the utility of aggression	CP-TF=-0.1, CP-BT=0.1, X^2 (1)=5.64, p=0.02

Table 35. Intensity of the strategy as a moderator of the effectiveness of the strategy: Summary of results (patient health and service utilization outcomes) (continued)

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Bickman et al., 2011[13] RCT/High RoB	G1: Weekly and cumulative 90-day feedback on patient symptoms and functioning to practitioners G2: Cumulative 90-day feedback on patient symptoms and functioning to practitioners only Single difference across arms frequency of quality monitoring mechanism (weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only) (comparison group strategy contained no components)	G1: 13 sites, 167 youths, 169 caregivers, 64 clinicians G2: 15 sites, 173 youths, 214 caregivers, 80 clinicians Total scales analyzed (breakdown by trial arm NR): youth, N=1,341; clinicians, N=1,291; caregivers, N=935	Youth Estimated coefficient of membership in feedback group at baseline Estimated coefficient of slope (time in weeks) Estimated coefficient of interaction of membership in feedback group and slope	0.02, SE: 0.10, p>.005 -0.001, SE: 0.002, p<.0001 -0.01, SE: 0.002, p<.001
Epstein et al., 2007[75]	G1a: Patients whose physicians did conduct a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during medication management program G 1b: Patients whose physicians did not conduct a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during medication management program G2: Control All (both) components differed across arms (comparison group includes neither of these components)	G1 (patients): 29 G1b (patients): 30 G2 (patients): 87	Mean scores for combined parent and teacher ratings of ADHD symptoms NR F score for strategy effect on combined parent and teacher ADHD ratings in subgroup of children who received a titration trial in G1 (compliers) Reduction in DSM-IV symptomatology	F4,124=3.80, p<0.01 G1a vs. G2: t114=−2.72, p=0.008, effect size=0.25 G1b vs. G2: t57= − 3.568, p=0.001, effect size=0.47

ADHD = attention deficit hyperactivity disorder; CFS =Contextualized Feedback Systems; CP-BT = Coping Power-Basic Training; CP-TF = Coping Power-Training plus Feedback; CI = confidence interval; *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV); G = group; N = number; NR = not reported; RCT = randomized controlled trial; RoB = risk of bias; SE = standard error.

The study on frequency of feedback to therapists found that effect sizes for symptoms and functional status increased by 50 percent for youth ratings of their own status, to 0.27, and by 66 percent for clinician reports, to 0.40 (p<0.001). The effect size did not increase for caregiver reports of adolescent functioning status.[13]

The study on using titration trials within a study of collaborative consultative model found that patients whose physicians conducted a titration trial had lower combined parent and teacher ratings of ADHD symptoms but did not have an effect on DSM-IV rated symptomatology.[75]

Strength of Evidence

We graded the single publication that presented data from a cluster RCT as having low strength of evidence for benefit that more intense treatment improved patient access to care and low strength of evidence for no benefit that more intensive treatment improved patient treatment engagement and practitioner protocol adherence/program fidelity in a strategy testing professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control (Table 36). We also graded this trial as having low strength of evidence for benefit that greater training intensity was associated with greater improvements in mental health symptoms (Table 37). We graded the other two publications that examined the moderating effect of training intensity as having insufficient strength of evidence for a strategy testing weekly and cumulative 90-day feedback on patient symptoms and functioning to practitioners versus cumulative 90-day feedback on patient symptoms and functioning to practitioners only on mental health symptoms and functional status and for a strategy testing practitioners who conducted a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during a medication management program versus practitioners who did not conduct a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during a medication management program versus control on mental health symptoms.

Table 36. Intensity of the strategy as a moderator of the effectiveness of the strategy: Detailed strength of evidence (intermediate outcomes)

Moderator and Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Training intensity: patient access to care (G1: Professional training plus feedback to implement an EBP intervention, G2: Professional training only to implement an EBP intervention, G3: Control)	1; 511 patients	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit More intensive training led to improved access-to-care ratings (sessions scheduled) for both children and for parents
Training intensity: patient treatment engagement (G1: Professional training plus feedback to implement an EBP intervention, G2: Professional training only to implement an EBP intervention, G3: Control)	1; 511 patients	High	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit More intensive training associated with no differences in child or parent sessions attended or parent ratings of treatment engagement, although treatment engagement for child rated higher for more intensive training group

Table 36. Intensity of the strategy as a moderator of the effectiveness of the strategy: Detailed strength of evidence (intermediate outcomes) (continued)

Moderator and Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Training intensity practitioner protocol adherence and program fidelity (G1: Professional training plus feedback to implement an EBP intervention, G2: Professional training only to implement an EBP intervention, G3: Control)	1; 511 patients	High	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit More intensive training did not lead to significant differences in mean numbers of strategy objectives completed, but did lead to increases in the numbers of contacts between practitioners and trainers in the CP-TF group

Table 37. Intensity of the strategy as a moderator of the effectiveness of the strategy: Detailed strength of evidence (patient health and service utilization outcomes)

Moderator and Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Training intensity: mental health symptoms (G1: Professional training plus feedback to implement an EBP intervention, G2: Professional training only to implement an EBP intervention, G3: Control)	1; 511 patients	High	Unknown (single study)	Direct	Precise	Undetected	Low for benefit More intensive training associated with greater improvements in mental health symptoms
Training intensity: mental health symptoms and functional status (G1: Weekly and cumulative 90-day feedback on patient symptoms and functioning to practitioners G2: Cumulative 90-day feedback on patient symptoms and functioning to practitioners only) t	1; N of practitioners unclear	High	Unknown (single study)	Direct	Unknown	Undetected	Insufficient Effect sizes for child and parent ratings of symptoms and functional status improved significantly in the more intensive training group but precision is unknown and study limitations are high

Table 37. Intensity of the strategy as a moderator of the effectiveness of the strategy: Detailed strength of evidence (patient health and service utilization outcomes) (continued)

Moderator and Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Training intensity: mental health symptoms (G1a: Patients whose physicians did conduct a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during medication management program G1b: Patients whose physicians did not conduct a titration trial as part of a collaborative consultative treatment service to promote the use of titration trials and periodic monitoring during medication management program G2: Control)	1; 144 patients	High	Unknown (single study)	Indirect	Unknown	Undetected	Insufficient Unknown precision and high study limitations

ADHD = attention deficit hyperactivity disorder; CP-TF = Coping Power-Training plus Feedback; N = number.

Moderating Effects of Process Characteristics—Fidelity to EBP

Study Description

One study[72] (medium risk of bias) studied a pay-for-performance strategy to improve treatment implementation of an EBP for adolescent substance use disorders by comparing a P4P condition to an implementation-as-usual control condition IAU. The EBP implemented to be used by all therapists was the A-CRA; each organization received standardized funding, training, and coaching from the treatment developers. After finding a significant association between target A-CRA and remission status (OR, 1.91; 95% CI, 1.02 to 3.58; p=0.04), the interaction between condition assignment and target A-CRA was examined with respect to patient remission status.

Patient Health and Service Utilization Outcomes

Fidelity to the EBP (i.e., meeting target A-CRA) did not significantly moderate the association between treatment group and patient remission status (authors do not report specific effect sizes other than p=0.37) (Table 38).

Table 38. Intensity of the strategy as a moderator of the effectiveness of the strategy: Summary of results

Study Design/Risk of Bias	Study Arms Differences in Strategy Components Across Study Arms	N Analyzed	Outcome Reported by Study and Time Period	Results
Garner et al., 2012[72] RCT/*RoB*	G1: Paying practitioners for performance (P4P) for successfully delivering of an EBP intervention (A-CRA) G2: Implementation of an EBP intervention (A-CRA) as usual	G1:14 organizations, 49 therapists, G2:15 organizations, 49 therapists	Patient health and service utilization outcomes: Mental health symptoms, syndromes, or disorders as measured by patient remission at 6-month followup	Fidelity did not moderate the association between treatment group and patient remission status (no effect sizes reported, p=0.37)

A-CRA= Adolescent Community Reinforcement Approach; EBP = evidence-based practice; G = group; RCT = randomized controlled trial; RoB = risk of bias; P4P = pay for performance.

Strength of Evidence

We graded a single publication that presented data from an RCT examining financial incentives provided to the practitioner for successful implementation of an EBT to treat adolescents with substance use disorders, of medium risk of bias, as having low strength of evidence that fidelity to the EBP did not moderate the effect of the strategy on patient remission (Table 39).

Table 39. Detailed strength of evidence for moderating effect of fidelity to EBP in P4P

Outcome	Number of Studies; Subjects	Study Limitations	Consistency	Directness	Precision	Reporting Bias	Strength of Evidence Grade Magnitude of Effect
Final outcome: mental health symptoms (G1: Paying practitioners for performance (P4P) for successfully delivering of an EBP intervention (A-CRA) G2: Implementation of an EBP intervention (A-CRA) as usual)	1; 600	Medium	Unknown (single study)	Direct	Precise	Undetected	Low for no benefit There was not a significant moderating effect of fidelity to EBP (meeting target A-CRA) on the association between treatment group and patient remission status (p=0.37)

A-CRA= Adolescent Community Reinforcement Approach; EBP = evidence-based practice; P4P = pay for performance.

Discussion

This chapter summarizes key findings and strength of evidence for each Key Question (KQ), followed by a summary of the limitations of the review, limitations of the evidence base, gaps in the evidence that may benefit from future research, and overall conclusions.

Key Findings and Strength of Evidence

Key Question 1. Effectiveness of Strategies to Improve Mental Health Care for Children and Adolescents

Overview

The strategies included in this review were heterogeneous and difficult to categorize. We encountered a large degree of uncertainty and overlap when classifying the examined strategies as implementation, dissemination, and quality improvement (QI) (our initial taxonomy). We then shifted to the EPOC taxonomy to identify individual components and groups of components. This taxonomy allowed us to group strategies in two categories: (1) professional training strategies with professional components only or (2) financial or organizational change strategies, with at least one financial or organizational component in addition to professional components. Most strategies were complex and included multiple (2 to7) different components.

Tables 40 and 41 present strength of evidence grades for professional training and financial or organizational change strategies, respectively. We graded 14 outcomes for professional training and 16 for financial or organizational change, and over half of these grades are insufficient or low for no benefit. Nonetheless, we found evidence that a majority of strategies had at least some evidence of effectiveness. Ten studies reported in 9 publications (i.e., 9 strategies) had at least one outcome rated as low for benefit. One study had a single outcome rated as moderate for benefit. Overall, 3 of 5 strategies (6 studies) reported in 5 publications classified as having Professional Only components and 7 of 9 strategies classified as having at least one Organizational or Financial component had at least one outcome rated as low or medium for benefit. Therapists in the pay-for-performance group were over twice as likely to demonstrate implementation competence compared with implementation-as-usual therapists.[72] Other outcomes for which we found low strength of evidence of benefit included improved practitioner adherence from training practitioners to monitor metabolic markers,[82] computer decision support plus EHR that included diagnosis and treatment guidelines,[78] and Internet portal providing practitioner access to practice guidelines;[74] improved practitioner morale, engagement, and stress from organizational change;[73] improved patient access to care, parent satisfaction, treatment engagement, and therapeutic alliance from training nurses to educate parents about EBPs;[71] improved child behavior problems in the short term (6 months) and out-of-home placements from organization change,[14] improved patient functional status from weekly feedback on patient symptoms and functioning to practitioners;[13] and improved service utilization from training practitioners on medication monitoring[82] and appropriate identification and referral of cases.[79]

Only four strategies (1 study each) consistently provided insufficient or evidence of no benefit across all reported outcomes. These included a strategies testing augmented active learning versus computerized routine versus routine practitioner workshop to implement an EBP,[76] collaborative consultation treatment service to promote the use of titration trials and

periodic monitoring during medication management versus control,[75] an Intensive Quality Assurance system versus workshop only to implement an EBP intervention,[80] and professional training plus feedback to implement an EBP intervention versus professional training only to implement an EBP intervention versus control (3 arms).[77]

We found no clear patterns of effectiveness associated with categorization of the strategy according to EPOC components. The studies varied with respect to the number and types of active components (that is, differences in components that comprised the treatment group strategy and the comparison group strategy). In some studies, the treatment group contained myriad components and the comparison group contained none of those components. In other studies, both the treatment and comparison groups tested strategies with multiple components, with varying number of differences in components across arms. When both arms receive several active strategies, the Hawthorne effect may explain lack of effectiveness. We did not find, however, any consistent patterns of effectiveness regarding the number of active components, that is, we did not find that studies that employed strategies with a single active component had any better or worse effect on outcomes than those that employed multiple active components.

Additional heterogeneity arose from several other sources. With the exception of the two studies reported in 1 publication[71] and two studies reporting variants of a similar strategy,[14, 73] none of the other studies tested similar strategies. The outcomes of the studies varied widely, as did the settings (community-based hospitals and clinics, general practice/primary care, home-based mental health systems, schools). Differences in the target of each strategy (e.g., practitioners, practices, systems) further precluded quantitative synthesis of our findings.

The absence of evidence on several factors of interest further limits our conclusions. We found no evidence of studies examining several intermediate outcomes, particularly system-level intermediate outcomes, as well as final patient health outcomes such as co-occurring conditions, mortality, or quality of life. We also found no evidence of strategies testing multiple components of the EPOC taxonomy, including any regulatory components and little evidence on strategies with financial components. We rated 6 studies as having unclear risk of bias and 6 studies reported in 5 publications as having high risk of bias. Thus, out of 15 studies included in our review, only 1 study had low risk of bias and 2 had medium risk of bias. Various issues with study design, attrition, and incomplete information reported by study authors precluded most of these studies from having a low or medium risk of bias.

The uncertain or high risk of bias of most of these studies affected the overall strength of evidence grades, as did the inclusion of single studies for each strategy examined.

Table 40. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: satisfaction/ acceptability 1 RCT; 115 therapists[76] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Insufficient for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, single imprecise measure, CIs cross the line of no difference)
Practitioner: adherence/ fidelity 1 RCT; 115 therapists[76] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Low for no benefit for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, multiple imprecise measures with CIs crossing the line of no difference)[a]
Practitioner: adherence/ fidelity 1 ecological study; practitioners of 2,376 patients[82] 38.3% of patients had MMT in the charts after program implementation; drop in the prevalence of SGA prescribing from 15.4% in the pre-MMTP period to 6.4% in the post-MMTP period (p<0.001)	Educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator	Low for benefit for patient medication monitoring training program for practitioners vs. usual care (high study limitations, precise outcomes)
Practitioner: competence/skills 1 RCT; 115 therapists[76] No statistically significant difference between groups	*Augmented routine professional training workshop with active learning component:* workshop with behavioral role play and small group activities *Computerized routine professional training workshop:* distribution of program's educational materials delivered via the computer *Routine professional training workshop:* workshop with didactic instruction	Low for no benefit for augmented active learning vs. computerized routine vs. routine professional training workshop to implement an EBP (low study limitations, multiple imprecise measures with CIs crossing the line of no difference)

Table 40. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient: access to care 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[71] Improved parent ratings of access to care (mean difference between groups ranged from 0.08 to 2.1 points in Study 1 and 0.6 to 1.9 in Study 2, scale 1-5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: access to care 1 RCT; 110 practices, 79 patients[79] RR of referral to early intervention: 1.20, 95% CI, 0.74 to 1.95, p=0.48	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient: satisfaction 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[71] Improved parent ratings of satisfaction of care by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: treatment engagement 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[71] Strategy improved parent ratings of treatment engagement by a mean of 0.9 in Study 1 and 2.5 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)
Patient: therapeutic alliance with provider 2; 172 parents/caregivers in Study 1 (quasi-experimental), 51 in Study 2 (RCT)[71] Strategy improved parent ratings of therapeutic alliance by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Distribution of educational materials, educational meetings, educational outreach visits, patient-mediated interventions	Low for benefit of protocol to train nurses to educate parents about EBPs vs. typical services (high study limitations, consistent, precise)

Table 40. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health or service utilization outcome: changes in mental health status 1 RCT; 511 patients[77] CP-TF and CP-BT groups had fewer behavioral problems as rated by teachers (beta=-0.41, SE=0.16, p=0.01). There were no significant differences in teacher ratings of behavioral problem for the CP-BT versus comparison group or for any comparisons of behavioral problems as rated by parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Patient health or service utilization outcomes: changes in mental health status 1; 158 patients for detainment under Mental Health Act, 83 patients for recovery[79] No statistically significant differences between groups	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient health or service utilization outcomes: socialization skills and behaviors 1 RCT; 511 patients[77] CP-TF had fewer minor assaults than comparison group as reported by the child (beta=-0.25, SE=0.12, p=0.03) and social/academic competence as reported by the teacher (beta=0.35, SE=0.13, p=0.01). These differences were not significant for the CP-BT versus comparison groups, nor were any significant differences found between groups on social skills as rated by parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Patient health or service utilization outcomes: service utilization 1 ecological study; 253 patients (82 before, 171 after MMTP implementation (SGA-treated subset of patients from overall N of 2,376)[82]	Educational materials, educational meetings, educational outreach visits, reminders, and online access and project coordinator	Low for benefit for patient medication monitoring training program for practitioners vs. usual care (high study limitations, precise results)

Table 40. Summary of results of the effectiveness of professional training strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health or service utilization outcomes: service utilization 1 RCT: 68 patients for number of consultations in primary care following the intervention and duration of untreated psychosis and delay in reaching early-intervention services[79] Patients in the professional training group averaged 223.8 days shorter for time from the first decision to seek care to the point of referral to an early-intervention service than patients in the control group	Educational meetings, local consensus process, educational outreach visits, marketing	Low for benefit for professional training to identify and refer cases vs. treatment as usual (high study limitations, precise results)

[a]We rated this outcome as low for no benefit rather than insufficient because of the consistency of results from multiple measures.

CI = confidence interval; CP-BT = Coping Power-basic training; CP-TF = Coping Power-training plus feedback; EBP = evidence-based practice; MMT= metabolic monitoring program; MMTP = metabolic monitoring training program; NR = not reported; RCT = randomized controlled trial; SGA = second generation antipsychotic; SE = standard error; vs. = versus.

Table 41. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: adherence/fidelity 1 RCT; 84 patients[78] Practitioner adherence improved through uptake of guidelines for diagnostic assessment (aOR, 8.0; 95% CI, 1.6 to 40.6); more reporting of 3 of 4 symptom domains at diagnosis No statistically significant differences on practitioner adherence through reassessment of symptoms at 3 months, adjustment of medications, mental health referral	Patient-mediated intervention, reminders, quality monitoring	Low for benefit for computer decision support plus Electronic Health Record (EHR) that included diagnosis and treatment guidelines vs. computer decision support plus EHR only on two of three measures (uptake of guidelines and measurement of ADHD symptoms at diagnosis; medium study limitations, imprecise results with small number of events, large magnitude of effect); insufficient for reassessment of symptoms at 3 months, adjustment of medications, and referral (medium study limitations, imprecise results [CI cross the line of no difference])
Practitioner: adherence/fidelity 1 RCT; 38 practitioners[75] Interaction for uptake of titration trials β=-0.283; SE, 0.09; p<0.01 Uptake of medication monitoring trials: p=NS, details NR	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])
Practitioner: adherence/fidelity 1 RCT; 746 patients[74] Strategy appeared to improve 4 of 5 examined outcomes that measured practitioner protocol adherence/program model fidelity outcomes (range mean change in outcome between groups range from 16.6 to -50) but estimates were very imprecise, with large confidence intervals	Distribution of educational materials, educational meetings, patient-mediated interventions, audit and feedback, reminders, provider incentives, quality monitoring and staff organization	Low for benefit for internet portal providing practitioner access to practice guidelines vs. wait-list control (medium study limitations, imprecise [wide CIs])

Table 41. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: adherence/fidelity 1 RCT; variable by analysis[14] No statistically significant difference between groups for caregiver-reported MST therapist, audio-coded ratings of therapist adherence, ratings of supervisor adherence	*Organizational change and an EBP intervention* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work, quality monitoring *EBP intervention only:* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, quality monitoring *Organizational change only:* Distribution of educational materials, educational meetings, educational outreach visits, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for no benefit for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, details on precision NR)
Practitioner: adherence/fidelity 1 RCT; N of practitioners unclear[13] Two-thirds of practitioners did not view Web modules	Frequency of quality monitoring mechanism	Insufficient for weekly and cumulative 90-day vs. cumulative 90-day feedback on patient symptoms and functioning to practitioners (high study limitations, unknown precision)
Practitioner: adherence/fidelity 1 CCT; 30 practitioners, N of caregiver and patient reports and monthly data points NR[80] Study does not provide sufficient detail to judge magnitude of effect	Patient incentives and quality monitoring	Insufficient for Intensive Quality Assurance system vs. workshop only to implement an EBP intervention (high study limitations, imprecise results)
Practitioner: morale, engagement, stress 1 RCT; 197 practitioners in 26 programs[73] Trends toward improving all domains, but nonoverlapping CIs for only some domains	Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for benefit for organizational change vs. control (medium study limitations, details on precision NR)

Table 41. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Practitioner: competence/skills 1 RCT; 38 practitioners[75] Lower odds with overlapping confidence intervals of citing obstacles in 6 of 8 measures (2 reach statistical significance)	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size]))
Practitioner: incentives 1 RCT; 49 therapists and 936 patients[72] Therapists in the P4P group were over twice as likely to demonstrate implementation competence compared with IAU therapists (Event Rate Ratio, 2.24; 95% CI, 1.12 to 4.48); patients in the P4P condition were over five times as likely to meet target implementation standards (i.e., to receive specific numbers of treatment procedures and sessions) as IAU patients (OR, 5.19; 95% CI, 1.53 to 17.62) but confidence intervals were wide	Provider incentives	Moderate for benefit for paying practitioners for performance of successful delivery of an EBP intervention vs. implementation as usual (medium study limitations, precise results)
Patient: access to care 1 CCT; 4 pediatric practices, 20,917 children with primary care visits[81] Improvement in patient access to care (attending first Triple P visit; OR, 3.10; 95% CI, 1.63 to 5.89)	Changes in scope and nature of benefits and services	Low for benefit for co-location of an EBP program in primary care vs. enhanced referral to an EBP program (high study limitations, precise results)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 144 patients[75] F score for decrease in combined parent and teacher ratings of ADHD symptoms for group x time interaction: $F_{2, 144} = 0.44$, p=0.65	Audit and feedback and clinical multidisciplinary teams	Insufficient for collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])

Table 41. Summary of results of the effectiveness of organizational or financial strategies to improve mental health care among children and adolescents (KQ 1) (continued)

Outcome category, outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 567 caregivers of youth[14] At 6 months, lower child behavior problem scores for the MST plus ARC arm, but not ARC or MST only compared with control At 18 months, no statistically significant difference in child behavior problem scores between groups 1 RCT; 615 youth[14] Lower rate of out-of-home placement for MST or ARC, but not a statistically significant added benefit of MST plus ARC compared with control (34%)	*Organizational change and an EBP intervention* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work, quality monitoring *EBP intervention only:* Distribution of educational materials, educational meetings, educational outreach visits, audit and feedback, quality monitoring *Organizational change only:* Distribution of educational materials, educational meetings, educational outreach visits, training and cognitive models to improve effectiveness, satisfaction of providers with conditions of their work	Low for benefit at 6 months and no benefit at 18 months for child behavior problem scores for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, precise results) Low for benefit for out-of-home placements for organization change and an EBP intervention vs. an EBP intervention only vs. organizational change only vs. control (medium study limitations, precise results)
Patient health and service utilization outcomes: changes in mental health status 1 RCT; 600[72] No statistically significant differences between groups	Provider incentives	Low for no benefit for paying practitioners for performance of successful delivery of an EBP intervention vs. implementation as usual (medium study limitations, precise results)
Patient health and service utilization outcomes: functional status 1 RCT; 340 youth,144 clinicians, 383 caregivers[13] Membership in the weekly feedback group increased the rate of decline in symptoms and functioning severity scale by 0.01	Frequency of quality monitoring mechanism	Low for benefit for weekly and cumulative 90-day vs. cumulative 90-day feedback on patient symptoms and functioning to practitioners (high study limitations, precise results)
Patient health and service utilization outcomes: service utilization 1; 84 patients[78] Calculated OR: 2.195; 95% CI, 0.909 to 5.303; p=0.081, reported p-value in study=0.054	Patient-mediated intervention, reminders, quality monitoring	Insufficient for computer decision support plus EHR that included diagnosis and treatment guidelines vs. computer decision support plus EHR only (medium study limitations, imprecise results [CI cross the line of no difference])

ADHD = attentional deficit hyperactivity disorder; aOR = adjusted odds ratio; ARC = Availability, Responsiveness and Continuity; CCT = controlled clinical trial; CI = confidence interval; EHR = electronic health record; EBP = evidence-based

practice; MST = multisystemic therapy; NR = not reported; IAU = implementation as usual; OR = odds ratio; P4P = pay for performance; RCT = randomized controlled trial; SE = standard error; vs. = versus.

Key Question 2. Harms Associated With Strategies to Improve Mental Health Care for Children and Adolescents

Only one study evaluated the harms associated with strategies to improve mental health care for children and adolescents (Table 42). We graded this study as having insufficient strength of evidence of harms associated with a professional training strategy to improve access to early intervention for adolescents and young adults with psychosis, although it reported no adverse events or between-group differences in false positive referrals to primary care. Of note, no other studies reported on any of the a priori patient, organization, or other types of provider-related harms (Table 43).

Table 42. Summary of evidence of harms associated with strategies to improve mental health care among children and adolescents (KQ 2)

Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Patient: adverse events 1 RCT; 110 practices, 79 patients[79] No adverse events reported.	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)
Patient: false-positive referrals 1 RCT; 110 practices, 79 patients[79] No between-group differences in false-positive referrals to primary care	Educational meetings, local consensus process, educational outreach visits, marketing	Insufficient for professional training to identify and refer cases vs. treatment as usual (high study limitations, imprecise results)

RCT=randomized controlled trial; vs. = versus.

Table 43. Evidence of a priori harms (KQ 2) outcomes found

A Priori Harms Outcomes	Evidence Identified in the Review
Patient	
Lower treatment engagement/increased dropouts	None
Negative impact on therapeutic relationship	None
Side effects of EBP incorporated into strategy (e.g., adverse events, suicidality)	1 study (adverse events, false-positive referrals to primary care)
Patient dissatisfaction with care	None
Provider	
Burnout/exhaustion	None
Turnover	None
Resistance to strategy	None
Organization	
Cost	None
Failure to sustain EBP	None
Resistance to change	None
Resistance to strategy	None

EBP = evidence-based practice; KQ = Key Question.

Key Question 3. Moderators of the Effectiveness of Strategies to Improve Mental Health Care for Children and Adolescents

Overall, four included studies examined moderators of the effectiveness of strategies to improve mental health care for children and adolescents (Table 44). Three examined whether training intensity influenced the degree of effectiveness, with two studies being graded as having insufficient strength of evidence. The third study that examined the moderating effect of training intensity had low strength of evidence for benefit of the moderating effects of training intensity on both patient intermediate (access to care) and patient health and service utilization outcomes (change in mental health status). School therapists receiving more intensive training had greater improvements in patient access to care ratings (sessions scheduled) for both children and for parents and greater improvements in mental health symptoms (i.e., less externalizing behaviors) than therapists receiving less intensive training. We were unable to combine the findings from these studies due to the heterogeneity in the strategies being tested. A fourth study examined the moderating effects of fidelity to the EBP (meeting target Adolescent Community Reinforcement Approach [A-CRA]) used as part of the strategy. We graded this study as having low evidence of no benefit for moderating the effect of the strategy, patient health outcome, and patient remission status. Also of note, we did not find studies that examined most of our a priori list of moderators such as patient characteristics, intervention characteristics other than training intensity, factors of the outer or inner setting/organizational factors, characteristics of involved individuals, process characteristics other than training fidelity, or other moderators such as length of followup (Table 45).

Table 44. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Patient: patient access to care 1 RCT; 110 practices, 79 patients[79] More intensive training led to improved access to care ratings (sessions scheduled) for both children and for parents	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Patient: treatment engagment 1 RCT; 110 practices, 79 patients[79] No significant differences between groups	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Practitioner: protocol adherence/program fidelity 1 RCT; 110 practices, 79 patients[79] No significant differences between groups	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for no benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)

Table 44. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3) (continued)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; 511 patients[77] More intensive training associated with greater improvements in mental health symptoms	*Professional training plus feedback*: educational training, educational meetings, educational outreach visits, marketing, and online access and project coordinator *Professional training only*: educational meetings and marketing	Low for benefit for moderating effect of training intensity on professional training plus feedback to implement an EBP intervention vs. professional training only to implement an EBP intervention vs. control (medium study limitations, precise results)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; N of practitioners unclear[13] Effect sizes for child and parent ratings of symptoms improved significantly in the more intensive training group	Weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only	Insufficient for moderating effect of training intensity on weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners (high study limitations, unknown precision)
Training Intensity Patient health and service utilization: mental health symptoms 1 RCT; 197 practitioners in 26 programs[73] Reduction in mental health symptoms in the compliers group was significantly greater than that seen in the control group (t(114)=-2.72, p=.008, effect size=0.25) and in the noncomplier group (t(57)= -3.568, p=.001, effect size=0.47).	Audit and feedback and clinical multidisciplinary teams	Insufficient for moderating effect of training intensity on collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control (high study limitations, imprecise results [small sample size])
Training Intensity Patient health and service utilization: functional status 1 RCT; 49 therapists and 936 patients[72] No significant moderating effect of fidelity to EBP (meeting target A-CRA) on the association between treatment group and patient remission status	Provider incentives	Low for no benefit for moderating effect of fidelity to EBPs on paying practitioners for performance in successfully delivering of an EBP intervention vs. implementation as usual (medium study limitations, precise results)

Table 44. Moderators of the effectiveness of strategies to improve mental health care among children and adolescents (KQ 3) (continued)

Moderator Outcome Category, Outcome Number of Studies; n of Individuals Results	Active Strategy Component	Strength of Evidence (Domain-Specific Ratings)
Training Intensity Patient health and service utilization: functional status 1 RCT; N of practitioners unclear[13] Effect sizes for child and parent ratings of functional status improved significantly in the more intensive training group	Weekly feedback to providers and cumulative 90 day feedback versus 90 day feedback only	Insufficient for moderating effect of training intensity on weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners (high study limitations, unknown precision)

A-CRA = Adolescent Community Reinforcement Approach; ADHD = attention deficit hyperactivity disorder; EBP = evidence-based practice; N/A = not available; RCT=randomized controlled trial; SOE =strength of evidence.

Table 45. Evidence of a Priori Moderators Found

A Priori Harms Outcomes	Evidence Identified in the Review
Patient characteristics (age, gender, cognitive functioning, diagnosis/severity of mental health problem, comorbid conditions, cotreatments, race/ethnicity)	None
Intervention characteristics (complexity, manualized or not, intensity/frequency/duration, adjustment of intervention to fit context)	3 studies (Intensity only)
Outer setting (external policy, incentives, availability of alternative care systems)	None
Inner setting/organizational factors (type of outpatient setting, structure/size, culture, implementation climate, readiness of organization for implementation)	None
Characteristics of involved individuals (provider type, knowledge, beliefs, self-efficacy, leadership, education, certifications, accreditation policies, standards, and years of practice)	None
Process characteristics (fidelity to the planned strategy, fidelity to the EBP, use of champions or supervision/oversight)	1 study (fidelity to the EBP only)
Other: length of followup	None

EBP = evidence-based practice.

Findings in Relationship to What Is Already Known

This systematic review contributes to the literature on QI, implementation and dissemination strategies targeting systems, and organizations or practitioners of mental health care to children and adolescents in several ways. First, this review offers an updated examination of the literature. Two recent systematic reviews have addressed this topic. A systematic review by Barwick et al.[38] examined 12 studies of knowledge translation interventions and strategies related to the delivery, organization, or receipt of child and youth mental health services that were published between 2001 and 2009.[38] All 12 studies reported significant changes in behaviors as a result of knowledge translation, although the quality of studies was limited by insufficient or unclear reporting and low sample sizes. The authors also noted that the behaviors were largely self-reported rather than observed, and several studies involved simulated situations rather than real-world settings. A systematic review by Novins et al. in 2013 examined studies of dissemination and implementation of mental health EBPs, including substance abuse, between 1991 and December 2011.[39] The authors reported that a majority of the included articles were observational rather than experimental and that the strongest empirical evidence existed for fidelity monitoring and supervision. Both prior reviews called for additional studies on these topics because of the dearth of sufficient evidence in this field.

Our review includes studies from inception through November 2014. The inclusion criteria for the current study are more narrowly defined than the aforementioned reviews in terms of study design and population. Barwick et al. and Novins et al. included studies that focused on teacher training for behavior change, in contrast to the current study focusing primarily on mental health practitioners.

Our study focused on strategies targeting a more narrowly defined mental health population. We excluded studies examining strategies focusing specifically on groups of children with developmental disabilities including autism, because of the heterogeneity in strategies used and types of systems involved in their care, in contrast to Barwick's review, which included studies in which the primary mental health population was diagnosed with autism. We also included only studies of youth with mental health symptoms and did not include prevention studies focused on populations not currently experiencing mental health symptoms. Novins et al. included several studies in a child welfare setting, where presumably many of the children are at risk for developing mental health symptoms, but the study was focused on preventive strategies. Barwick et al. excluded studies of children with substance abuse. Like Novins et al., we included studies of youth with substance abuse as the primary diagnosis. Another difference in the targeted mental health population is the age range. Barwick et al. defined youth more broadly, including studies with youth up to age 24 years; our study focused on studies examining sample made primarily of children through age 18 years.

We attempted to include only studies for which we could distinguish the effects of the strategy of interest from the underlying EBP and 11 of the 15 studies identified were RCTs. Two studies were CCTs, one was an ecological aggregated design, and one was quasi-experimental. Barwick et al. included 4 quasi-experimental designs, and Novins et al.'s review was much broader in scope, including descriptive and qualitative studies. Nonetheless, the two prior reviews and the current review each concluded, similarly, that there appears to be some evidence of efficacy of these strategies, but the field is too new to make definitive conclusions and additional research that includes well-designed studies with good reporting of methods is needed.

In addition to reviewing implementation strategies used to adopt and integrate EBPs into routine care and dissemination strategies used to disseminate evidence through increasing access to EBPs or people's motivation or ability to use and apply EBPs, our review also examined QI strategies, which were not explicitly included in the other reviews. Our study also highlighted the overlap between the definition of QI and dissemination/implementation studies.

Unlike other reviews, our review attempted to understand the moderating effects of different variables on effectiveness or harms. Moderators of interest included patient characteristics such as age, gender, race, cognitive ability, diagnosis, severity, coexisting conditions and cotreatments and intervention characteristics, such as complexity, manualized or not, intensity, frequency, or duration. Our review found little evidence available, emphasizing the need for future studies to examine these variables.

This review offered the opportunity to examine the current literature on QI, dissemination, and implementation strategies to improve mental health in children, seeking strategies that will assist in closing the gap between research evidence and practice. The results of this review suggest that additional research is needed to determine the best strategies for improved quality of care and the dissemination and implementation of QI in mental health care for children and adolescents. Numerous well-designed clinical trials of mental health interventions for youth exist, but our knowledge of how to best disseminate and implement these interventions remains limited because of (1) the limited number of studies conducted to date, (2) high risk of bias in the

studies we identified, and (3) low strength of evidence. Only one study was rated low risk of bias, 6 were rated as unclear, 2 had medium, and 6 studies in 5 publications had high. Studies were found to have high or unclear risk of bias due to high attrition rates, nonrandom assignment or lack of details about randomization method, no intention to treat models, failure to adjust for baseline differences, or failure to report on whether there were baseline group differences.

Our review highlights the fact that we still do not have adequate knowledge of best methods to transport EBPs to clinical settings. A better understanding of variables that can impede the dissemination and implementation of EBPs would be helpful to facilitate further development of QI and dissemination and implementation strategies for mental health interventions in youth. Once these factors are better understood, then strategies designed to address these variables can be developed and investigated through clinical trials. For example, Chorpita et al.[88] point out the need to address challenges faced by clinical providers such as concerns about how an EBP might address comorbidity, because much of clinical practice occurs in the generalist setting where the typical presenting patient is characterized by more than one type of problem or diagnosis, and the provider's perception of whether the treatment will be effective can be an important variable. They also note it is important to better understand the social processes relevant to dissemination, arguing that the social influence process is at least in part responsible for the implementation of different mental health treatments for children. They note that understanding more about the issue of training procedures as well as trainer characteristics is necessary to understanding how to change therapist practices, and understanding trainee characteristics is necessary to understand barriers to clinicians' use of treatment strategies. Considering variables such as these will provide a framework to guide future QI and dissemination/implementation strategy development.

As new quality improvement, dissemination, and implementation strategies are developed, we hope to see future clinical trials with more rigorous experimental designs. Future research efforts should focus on targeting QI, dissemination, and implementation strategies of EBPs that vary by provider characteristics, population characteristics, and setting characteristics. We found that future strategies are needed both for the QI, dissemination, and implementation of EBPs in psychotherapy treatments as well as medication treatments of mental illness in youth. Other important targets include the development of dissemination strategies for introducing mental health care into areas lacking in mental health care, for example, very rural areas with fewer mental health providers. In these areas especially, targeting primary care providers may be essential.

Applicability

Population

The studies in this review were focused on children with mental health and substance abuse problems. Developmental disorders such as autism and learning disabilities were excluded because they are often treated through different service systems than child mental health. Most studies were focused on mental health disorders (13 studies), with two additional studies focused on substance use disorders. Providers of the target strategies were practitioners with professional training such as psychiatrists, psychologists, and nurses. Studies focused on strategies delivered by nonprofessionals such as teachers were excluded.

The age range of children included in the review was 2 years to 18 years. In addition, two studies focused on psychosis also included young adults because this is the age of first incidence of psychosis in most cases. The applicability of findings is therefore limited to professionally

trained practitioners of children and adolescents with mental health and/or substance use disorders.

Interventions

This review included dissemination, implementation, and QI strategies delivered by practitioners in typical service settings. All strategies reviewed were focused at the practitioner (e.g., training practitioners) or system (e.g., implementing a new medical management system) level. All of the studies included were carried out in real-world settings, in keeping with the topic of the review. The findings are applicable to service systems and practitioners who treat children and adolescents with mental health and/or substance use disorders.

Comparators

Comparison conditions included usual treatment, lower-intensity versions of the strategy under study, and pre-strategy implementation cases in one study implementing a system-level strategy within a hospital.[71]

Outcomes

Outcomes examined in the studies included intermediate practitioner and intermediate patient, but not intermediate system, outcomes. Thus, no studies examined intermediate system outcomes such as feasibility, update, timeliness, penetration, sustainability, and resources, including costs. In studies not implementing EBP interventions, patient health or service utilization outcomes were required to ensure that the strategy had an effect on ultimate outcomes. Several patient health outcomes of interest such as comorbidity and mortality were not examined in any included studies.

In addition, only one study examined a single harm, patient side effects, associated with a tested strategy. Finally, only four studies examined two moderators of interest (training intensity and fidelity to protocol), so the findings of this review do not provide information about the other moderators of interest such as patient characteristics, outer or inner setting factors, characteristics of involved individuals, or length of follow-up.

Setting

The review included studies set in schools, mental health clinics, and primary care. KQ 1 results did not vary noticeably across settings. KQ 2 findings on the risk for improper referral in primary care[89] may be more applicable for this specific setting, but more studies are needed. One study measured the impact of colocation of behavioral health care within a larger health-based care system on access to evidence-based care. The findings suggest that such colocation may increase access to effective child mental health services. For many health care systems, this may be an achievable, structural change that, once in place, could have lasting impact. However, again, more studies are needed to replicate this finding before any generalizations can be made.

Implications for Clinical and Policy Decisionmaking

Our review found that the body of evidence is relatively recent and small and lacks depth in any single clinical area or intervention strategy within child mental health. In this review, across included studies, the strength of evidence for any intermediate outcome was rated moderate in one instance and low or insufficient for the remainder. We did not find moderate strength of evidence for any patient health and service utilization outcomes (system or patient level). The

lack of strong evidence needs to be interpreted in the context of the general development of research in pediatric care. Even in arguably more mature QI research areas such as pediatric intensive care, infectious disease, and pulmonary disease, QI strategies can change provider behavior but system- and patient-level changes may be difficult to demonstrate. For example, Okelo et al. published a large review of 68 QI strategy studies that all aimed to improve provider adherence to evidence-based asthma protocols. The review authors found notably more evidence for changing process outcomes compared with evidence for changing patient-level outcomes.[90] Furthermore, mental health research and pediatric mental health research have notably greater challenges than the rest of health care for a variety of complex reasons (e.g., challenges in studying children, difficulty studying brain and behavior, stigma, funding); it follows that QI research would be behind in depth and quality. Given the paucity of evidence for children, a question that arises is whether QI, dissemination, or implementation strategies tested in the adult literature might apply to children. However, there is literature suggesting that the generation of quality measures is weak for adults as well.[91] Those deciding research and funding priorities may consider the value of the lessons from more mature areas of QI research when making decisions on QI research in child mental health.

The meager findings here regarding QI research may also speak to the widely recognized gap between established EBPs in mental health and the limited number of practitioners or organizations actually providing those EBPs. For example, CBT is the widely accepted EBP for childhood anxiety. Of note, a recent Cochrane review of cognitive behavioral therapy (CBT) for childhood anxiety included 1,955 subjects from 41 studies.[92] Meanwhile, in the community, a minority of those with an anxiety disorder actually receive any treatment,[93] let alone CBT specifically.[94] One constraint to this field may simply be that an insufficient body of clinicians and organizations delivers EBPs to drive research. One clinical and policy implication, then, is the need for basic dissemination and infrastructure development for the delivery of EBPs. QI research in this space could then grow with a wide expansion of EBP availability, guiding future resource allocation.

QI concepts in child health and child mental health are becoming fixtures of modern clinical practice and continuing professional education, despite the shortcomings of evidence in the area. For example, the American Board of Pediatrics requires its members to demonstrate periodic participation in QI exercises with their own patients. Clinicians in child mental health are increasingly working in or for accountable care organizations. These practitioners and these organizations will be increasingly anxious for guidance on how to use evidence-based practices in an efficient and effective manner. If accountable care entities require evidence of high quality care in child mental health, a robust QI evidence base will be needed. Such a QI evidence base will need to be developed against a backdrop of health systems that provide already established EBPs far more than they currently do.

Limitations of the Systematic Review Process

Challenges in this systematic review arose with (1) defining the strategy of interest, (2) constructing the search strategy, and (3) applying prespecified inclusion/exclusion criteria.

Regarding defining the strategy of interest, although we identified pragmatic definitions of QI, implementation, and dissemination that had been applied previously in systematic reviews, we found their application to this review to be difficult. The lack of consistency in the terminology used in the published literature meant that the use of self-selected descriptors such as “QI,” “implementation,” or “dissemination” by study authors did not conform to our a priori

definitions of these types of studies or to the other similarly labelled studies in the field. As a result, we used EPOC taxonomy to characterize strategies by their primary focus of their components. As noted previously, we required multiple reviews of each included article and, on one occasion, outreach to authors to ensure that we interpreted the study correctly.

Regarding searches, we ran multiple iterations over a period of 7 months. We initially mirrored the search strategy in a previously published review[39] but had to make substantial changes to capture concepts or terms that were not indexed by the National Library of Medicine's Medical Subject Headings (MeSH).

Regarding the application of prespecified inclusion and exclusion criteria, we found that attempts to specify the population criteria to ensure greater homogeneity of included strategys resulted in challenges when reviewing the evidence. For example, our criterion that the system or clinic care for children and adolescents with *existing* mental health issues (rather than the *risk of* mental health issues only) was difficult to apply in some cases. To limit the inadvertent loss of relevant articles in a field with inconsistent use of terminology and inadequate indexing, we did not automatically exclude prevention studies in our searches. As a consequence, we found numerous studies with inadequate reporting that required judgment on whether the system addressed children and adolescents *at risk of* or *actually experiencing* mental health problems. For example, we encountered studies of adolescents in juvenile drug courts. Although the authors did not specify what proportion of adolescents experienced mental health issues, we relied on the clinical and substantive expertise of the team to judge that adolescents in juvenile drug courts would likely have substance abuse issues or externalizing behavior problems.

We included a broad range of eligible comparators in our protocol (usual care, or any other QI, implementation, or dissemination strategy). In reviewing full-text studies, we encountered otherwise eligible studies in which the intervention combined both a patient-level intervention and a system-level strategy to implement or disseminate that intervention. In such cases, the use of a usual care arm did not permit the authors to draw conclusions about the effect of the implementation or dissemination strategy apart from the underlying intervention.[61-67, 95]

Our inclusion and exclusion criteria were designed to capture QI, implementation, and dissemination studies across a range of strategies. As a consequence, they did not capture all variants of a particular strategy, along the timeline of its development and application in practice. Reviews focusing on a single strategy can evaluate variants of that strategy without using stringent criteria and construct a qualitative narrative on the arc of the development and implementation of that strategy, but they cannot speak to QI, implementation, or dissemination in general. In other words, we traded a depth of understanding on the development and application of individual strategies for breadth in included strategy types.

Limitations of the Evidence Base

We found relatively few studies of effectiveness of strategies to improve the mental health care of children and adolescents, although there is evidence that some are effective in improving both intermediate and patient health and resource utilization outcomes. We did not find any studies that focused on system-level intermediate outcomes, including the costs of these strategies.

The lack of a common language to describe even a basic concern such as the primary purpose of the strategies (dissemination, implementation, or QI) served as a hindrance to synthesis. Strategies varied significantly in the number of components; the reporting on these components was not always clear enough to adequately describe the strategy or fully understand the relative

importance of component parts. Studies often offered limited descriptions of "usual care" arms when compared with descriptions of experimental arms. Even with limited reporting, we found wide differences in the number, intensity, and differences in services offered in "usual care" arms. These differences sharply limit our ability to make statements about the overall effectiveness of these strategies as a class.

Only one study examined harms, and three studies examined moderators of strategy effectiveness Although the field generally acknowledges the vast array of potentially influential moderators in implementation research,[32] we uncovered only four studies on two moderators (intensity and fidelity). The paucity of evidence on issues such as fidelity and adaptation further limits our understanding of the minimum change in strategy needed to achieve a significantly different process or health outcome.

We rated several outcomes as insufficient or low strength of evidence because of the underlying heterogeneity or limited number of studies on specific strategy types, system or practitioner targets, or child or adolescent conditions. In some instances, our grades were limited by high risk of bias in included studies (arising from high attrition rates, failure to adjust analyses for baseline levels of key outcomes or clustering within practitioners or clinics, and failure to account for missing data).

Our ability to derive firm conclusions on the effectiveness of included strategies was also hindered by methodological and reporting issues in the literature. The strategies of relevance to this review generally required that study arms be defined at a systems level for pragmatic reasons (e.g., the intervention changes a system characteristic) or to avoid contamination (e.g., a provider exposed to a new QI strategy may have difficulty applying the strategy selectively to some patients). Observational studies in general are constrained in their ability to make causal assertions because of the risk of confounding; observational studies of systems interventions have an additional burden of accounting for secular and unmeasured financial or organizational changes that may influence outcomes. Cluster randomized trials with clearly specified protocols allow interventions to be allocated appropriately at the systems level, with the potential to avoid the constraints of confounding and unmeasured cointerventions. However, the analyses of results from these trials required controls for clustering. These analyses (requiring hierarchical linear modeling) were complex but were often not reported well enough in our included studies to permit an independent evaluation of the effect size,[77] precision of the effect,[73, 75, 78, 79] or risk of bias[77, 78] (Table 46). QI, implementation, and dissemination trials often fail to report on basic elements of study design and conduct, such as sequence generation, allocation concealment, fidelity to the intervention, and the risk of contamination or crossover. As suggested by Table 46, these lacunae in reporting occurred in all study design types, regardless of their underlying rigor and complexity.

Table 46. Studies with insufficient reporting on risk of bias domains

Domain with Insufficient Reporting to Assess Risk of Bias	Number of Studies
Eligibility criteria	2 nonrandomized studies,[80, 81] 1 trial[78]
Randomization sequence generation	7 trials[13, 14, 71, 75, 77, 78]
Allocation concealment	8 trials[13, 14, 71, 73, 75, 77, 78]
Similarity of baseline characteristics	6 trials[73-78]
Fidelity to intervention	4 nonrandomized studies,[71, 80-82], 6 trials[13, 14, 71, 77-79]
Overall attrition	2 nonrandomized studies[80, 81]
Attrition by study arm (differential attrition)	2 nonrandomized studies,[80, 81] 4 trials[73, 76, 77, 79]
Risk of contamination or crossover	4 nonrandomized studies,[71, 80-82] 10 trials[13, 14, 71, 72, 74-79]

Research Recommendations

The evidence base is marked by a small number of studies on heterogeneous strategiesand outcomes focusing on intermediate, health, and resource utilization outcomes and very few studies on harms or moderators. Some additional research is forthcoming: we found three ongoing trials in a review of clinicaltrials.gov that may be applicable to future reviews on the topic (NCT02097355, NCT01829308, NCT02271386). The first identified trial examines the impact of a Web-based patient management and monitoring system (Integrated Clinical Information Sharing System (ICISS)) that was designed to track patients' disease symptoms and response to therapy over time in pediatric patients with ADHD, asthma, autism, depression, and/or epilepsy (NCT02097355). The second identified trial examines the implementation of two evidence-based intervention strategies of SBIRT (e.g., generalist versus specialist) on the outcomes of adolescent alcohol, tobacco, other drug use, and HIV risk behaviors (NCT01829308). The third identified trial examines the impact of educating and supporting primary care providers in the implementation of EBPs for ADHD (NCT02271386). All three studies will fill some of the evidence gaps by providing additional quantitative data on the effectiveness of strategies to improve intermediate and patient centered outcomes. We outline specific suggestions for future studies below.

Third-party payers are paying increasing attention to quality metrics, as health care systems move to accountable care models. We found no studies on regulatory components and just one study testing a financial component.[74] In addition to expanding the modest body of evidence thus far on professional training and financial or organizational change strategies, new studies should additionally evaluate regulatory and financial components to support the needs of accountable care organizations in the near future.

We did not find evidence on the majority of the outcomes that we specified a priori. Of particular note is that six strategies relied on EBPs and therefore did not report patient health outcomes.[73, 74, 76, 80, 81, 83] In instances where fidelity to the original intervention is maintained, the assumption that the same level of effectiveness will occur in a new trial is reasonable and leads to an efficient use of research funds. Unfortunately, not all measured fidelity adequately. New strategies relying on EBPs must, at a minimum, report on fidelity so practitioners and policymakers can judge whether the strategy is, in fact, a new intervention, rather than implementation or dissemination of an existing intervention.

The risk of crossover or contamination is of particular concern in systems strategies, but only one study explicitly provided information on the risk of crossover or contamination. As noted earlier, very few studies offered information on fidelity or on unanticipated changes. Information on pragmatic issues around implementation (fidelity, adaptation, and minimum elements necessary to achieve change) may not necessarily require new studies on strategies with existing information; support of analyses from existing studies may fill some of the gap.

Although the failure to use EBPs results in a gap between potential and achieved outcomes, closing the gap requires more than just the use of an array of EBPs. What continues to be unknown is how to bridge the gap in the context of the finite resource of time allocated for a patient encounter. As expectations for documenting or checking off quality metrics for each action within a patient encounter increase, the risk of errors of omission or commission increase. For new information to be actionable, more evidence is needed on the relative merits of each action or strategy.

Future research in this area requires appropriately timed outcome measures. One potential explanation for the lack of consistent demonstration of effectiveness across the included studies could be that studies reported on outcomes too early, before strategies had a chance to take effect. The included studies generally measured systems outcomes over the course of the intervention. One study measured adherence to CBT after 3 months of consultation,[76] and a second measured referrals to early intervention services at 4 months after intervention.[79] Although a third trial measured outcomes at 6 months from baseline, the intervention was ongoing for some portion of that period.[74] Studies generally measured patient outcomes within 6 months of completion of the trial, with two exceptions that measured outcomes at 18[14] and 24 months,[77] respectively. The majority of included studies appropriately used cluster RCTs. Cluster RCTs, like pragmatic trials, need more resources and are harder to complete than conventional studies. An additional consideration of cluster RCTs relates to reporting. The studies we found were marked by poor reporting. Concerns about the inadequacies of reporting have been noted elsewhere in the literature.[96, 97] A recent tool, the StaRI, offers standards for reporting implementation studies that, if adopted widely, can significantly improve the utility of these studies and the pace of translation of evidence into practice.[98]

Conclusions

Because we generally found a single study on each strategy, the evidence does not permit us to have a high degree of confidence about the efficacy of any one strategy. Nonetheless, our findings may have relevance for policymakers who do not require a high level of casual certainty (for example, in commissioning pilot studies). Although we found insufficient or low strength of evidence of no benefit for the majority of outcomes that we evaluated, our findings suggest that several strategies can improve both intermediate and final health and resource utilization outcomes. Eleven studies reported in ten publications of the 15 included studies (10 publications) resulted in significant improvements in at least one examined intermediate or patient health or service utilization outcome. Moderate strength of evidence (from 1 RCT) supports pay-for-performance to improve implementation competence.[72] Low strength of evidence supports training practitioners to monitor metabolic markers and service utilization (1 observational study);[82] providing treatment guidelines through computer decision support (1 RCT)[78] or an Internet portal (1 RCT)[74] to improve practitioner adherence; changing organizational structures to improve practitioner morale, engagement, and stress (1 study)[73] and child behavior problems in the short term (6 months) and out-of-home placements (1 study);[14] training nurses to educate parents about EBPs to improve patient access to care, parent satisfaction, treatment engagement, and therapeutic alliance (1 RCT, 1 quasi-experimental study);[71] providing weekly feedback to practitioners on patient status to improve patient functioning (1 RCT);[13] and appropriately identifying and referring cases to improve utilization (1 RCT).[79]

In addition to differences in strategies tested and specific components of each strategy, heterogeneity in clinical conditions of targeted children and adolescents, practitioner types, and settings precluded definitive conclusions about the effectiveness of any one particular strategy. We were unable to judge the potential for harms associated with these strategies that may mitigate benefits based on the single included study with information on harms. The available evidence from four studies on two moderators does not permit us to make general conclusions about the conditions under which these strategies might work optimally.

References

1. O'Connell ME, Boat T, Warner KE, eds. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. National Academies Press: Washington, DC; 2009.

2. Perou R, Bitsko RH, Blumberg SJ, et al. Mental health surveillance among children--United States, 2005-2011. MMWR Surveill Summ. 2013 May 17;62 Suppl 2:1-35. PMID: 23677130.

3. Weisz JR, Weiss B, Han SS, et al. Effects of psychotherapy with children and adolescents revisited: a meta-analysis of treatment outcome studies. Psychol Bull. 1995 May;117(3):450-68. PMID: 7777649.

4. Dishion TJ, McCord J, Poulin F. When interventions harm. Peer groups and problem behavior. Am Psychol. 1999 Sep;54(9):755-64. PMID: 10510665.

5. Silverman WK, Hinshaw SP. The second special issue on evidence-based psychosocial treatments for children and adolescents: a 10-year update. J Clin Child Adolesc Psychol. 2008 Jan-Mar;37(1):1-7.

6. Chorpita BF, Daleiden EL, Ebesutani C, et al. Evidence-based treatments for children and adolescents: an updated review of indicators of efficacy and effectiveness. Clin Psychol-Sci Pr. 2011 Jun;18(2):154-72. PMID: WOS:000291755900009.

7. Hoagwood K, Burns BJ, Kiser L, et al. Evidence-based practice in child and adolescent mental health services. Psychiatr Serv. 2001 Sep;52(9):1179-89. PMID: 11533391.

8. Jensen PS, Goldman E, Offord D, et al. Overlooked and underserved: "action signs" for identifying children with unmet mental health needs. Pediatrics. 2011 Nov;128(5):970-9. PMID: 22025589.

9. Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: variation by ethnicity and insurance status. Am J Psychiatry. 2002 Sep;159(9):1548-55. PMID: 12202276.

10. Saloner B, Carson N, Cook BL. Episodes of mental health treatment among a nationally representative sample of children and adolescents. Med Care Res Rev. 2014 Jun;71(3):261-79. PMID: 24399817.

11. Satcher D. U.S. Public Health Service, Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Washington, DC: Department of Health and Human Services; Stock No. 017-024-01659-4 2008.

12. Williams J, Klinepeter K, Palmes G, et al. Diagnosis and treatment of behavioral health disorders in pediatric practice. Pediatrics. 2004 Sep;114(3):601-6. PMID: 15342827.

13. Bickman L, Kelley SD, Breda C, et al. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. Psychiatr Serv. 2011 Dec;62(12):1423-9. PMID: 22193788.

14. Glisson C, Schoenwald SK, Hemmelgarn A, et al. Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. J Consult Clin Psychol. 2010 Aug;78(4):537-50. PMID: 20658810.

15. Balas EA, Boron SA. Managing clinical knowledege for healthcare improvements. Stuttgart, Germany: Schattauer Verlagsgesellschaft2000.

16. Shojania KG, McDonald KM, Wachter RM, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies, Vol. 1: Series Overview and Methodology. Technical Review 9 (Contract No. 290-02-0017 to the Stanford University-UCSF Evidence-based Practices Center). AHRQ Publication No. 04-0051-1. Rockville, MD: Agency for Healthcare Research and Quality; Aug 2004.

17. Institute of Medicine. Crossing the Quality Chasm. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press2001.

18. McPheeters ML, Kripalani S, Peterson NB, et al. Quality Improvement Interventions To Address Health Disparities. Closing the Quality Gap: Revisiting the State of the Science. Evidence Report No. 208. (Prepared by the Vanderbilt University Evidence-based Practice Center under Contract No. 290-2007-10065.) AHRQ Publication No. 12-E009-EF. Rockville, MD: Agency for Healthcare Research and Quality; August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm

19. Dorsey S, Pullmann MD, Deblinger E, et al. Improving practice in community-based settings: a randomized trial of supervision - study protocol. Implement Sci. 2013;8:89. PMID: 23937766.

20. Schoenwald SK, Sheidow AJ, Letourneau EJ. Toward effective quality assurance in evidence-based practice: links between expert consultation, therapist fidelity, and child outcomes. J Clin Child Adolesc Psychol. 2004 Mar;33(1):94-104. PMID: 15028545.

21. McKibbon KA, Lokker C, Wilczynski NL, et al. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? Implement Sci. 2010 Feb 12;5(1):16. PMID: 21080976.

22. McCormack L, Sheridan S, Lewis M, et al. Communication and Dissemination Strategies to Facilitate the Use of Health-Related Evidence. Evidence Report/Technology Assessment No. 213. (Prepared by the RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13(14)-E003-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm

23. Rabin BA, Brownson RC, Haire-Joshu D, et al. A glossary for dissemination and implementation research in health. J Public Health Manag Pract. 2008 Mar-Apr;14(2):117-23. PMID: 18287916.

24. Brownson RC, Colditz GA, Proctor EK. Dissemination and Implementation Research in Health. New York, NY: Oxford University Press2012.

25. Lomas J. Diffusion, dissemination, and implementation: who should do what? Ann N Y Acad Sci. 1993 Dec 31;703:226-35; discussion 35-7. PMID: 8192299.

26. Proctor EK, Landsverk J, Aarons G, et al. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health. 2009 Jan;36(1):24-34. PMID: 19104929.

27. Aarons GA, Glisson C, Green PD, et al. The organizational social context of mental health services and clinician attitudes toward evidence-based practice: a United States national study. Implement Sci. 2012;7:56. PMID: 22726759.

28. Aarons GA. Mental health provider attitudes toward adoption of evidence-based practice: the Evidence-Based Practice Attitude Scale (EBPAS). Ment Health Serv Res. 2004 Jun;6(2):61-74. PMID: 15224451.

29. Glisson C, Landsverk J, Schoenwald S, et al. Assessing the organizational social context (OSC) of mental health services: implications for research and practice. Adm Policy Ment Health. 2008 Mar;35(1-2):98-113. PMID: 18085434.

30. Mendel P, Meredith LS, Schoenbaum M, et al. Interventions in organizational and community context: a framework for building evidence on dissemination and implementation in health services research. Adm Policy Ment Health. 2008 Mar;35(1-2):21-37. PMID: 17990095.

31. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health. 2011 Jan;38(1):4-23. PMID: 21197565.

32. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing

implementation science. Implement Sci. 2009;4:50. PMID: 19664226.

33. Meyers DC, Durlak JA, Wandersman A. The quality implementation framework: a synthesis of critical steps in the implementation process. Am J Community Psychol. 2012 Dec;50(3-4):462-80. PMID: 22644083.

34. Kaplan HC, Provost LP, Froehle CM, et al. The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. BMJ Qual Saf. 2012 Jan;21(1):13-20. PMID: 21835762.

35. Ringeisen H, Henderson K, Hoagwood K. Context matters: schools and the "research to practice gap" in children's mental health. School Psych Rev. 2003;32(2):153-68.

36. Rojas Smith L, Ashok M, Dy SM, et al. Contextual Frameworks for Research on the Implementation of Complex System Interventions. Methods Research Report. (Prepared by the RTI International-University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 14-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2014. www.effectivehealthcare.ahrq.gov/reports/final.cfm

37. Hoagwood K. Personal communication. 2014.

38. Barwick MA, Schachter HM, Bennett LM, et al. Knowledge translation efforts in child and youth mental health: a systematic review. J Evid Based Soc Work. 2012;9(4):369-95. PMID: 22830938.

39. Novins DK, Green AE, Legha RK, et al. Dissemination and implementation of evidence-based practices for child and adolescent mental health: a systematic review. J Am Acad Child Adolesc Psychiatry. 2013 Oct;52(10):1009-25 e18. PMID: 24074468.

40. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339(4):b2535. PMID: 19622551.

41. United Nations Development Program. Human development report. 2014. http://hdr.undp.org/en/content/human-development-report-2014. Accessed 29 May 2015.

42. Powell BJ, McMillen JC, Proctor EK, et al. A compilation of strategies for implementing clinical innovations in health and mental health. Med Care Res Rev. 2012 Apr;69(2):123-57. PMID: 22203646.

43. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy. Oslo: Norwegian Knowledge Centre for the Health Services2002. https://epoc.cochrane.org/epoc-taxonomy.

44. White CM, Ip S, McPheeters M, et al. Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews [posted September 2009]. Rockville, MD; 2009. http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318.

45. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. AHRQ Publication No. 12-EHC047-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2012. www.effectivehealthcare.ahrq.gov/

46. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928. PMID: 22008217.

47. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012 Feb;65(2):163-78. PMID: 21959223.

48. Sterne JAC, Higgins JPT, Reeves BC, et al. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI), Version 1.0.0. 24 September 2014, 2014. http://www.riskofbias.info. Accessed 29 May 2015.

49. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med. 2014 Feb 18;160(4):267-70. PMID: 24727843.

50. West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity. Methods Research Paper. AHRQ Publication No. 10-EHC070-EF. Rockville, MD: Agency for Healthcare Research and Quality; Sep 2010. http://effectivehealthcare.ahrq.gov/

51. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290- 2007-10056-I). AHRQ Publication No. 13(14)-EHC130-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm

52. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2014 Dec 20PMID: 25721570.

53. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011 Dec;64(12):1283-93. PMID: 21839614.

54. Bhandari M, Sprague S, Hanson B, et al. Health-related quality of life following operative treatment of unstable ankle fractures: a prospective observational study. J Orthop Trauma. 2004 Jul;18(6):338-45. PMID: 15213498.

55. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions. Methods Guide for Comparative Effectiveness Reviews. AHRQ Publication No. 11-EHC019-EF. Agency for Healthcare Research and Quality; January 2011. http://effectivehealthcare.ahrq.gov/

56. Bjorknes R, Manger T. Can parent training alter parent practice and reduce conduct problems in ethnic minority children? A randomized controlled trial. Prev Sci. 2013 Feb;14(1):52-63. PMID: 23135877.

57. Kjobli J, Hukkelberg S, Ogden T. A randomized trial of group parent training: reducing child conduct problems in real-world settings. Behav Res Ther. 2013 Mar;51(3):113-21. PMID: 23318242.

58. Grupp-Phelan J, McGuire L, Husky MM, et al. A randomized controlled trial to engage in care of adolescent emergency department patients with mental health problems that increase suicide risk. Pediatr Emerg Care. 2012 Dec;28(12):1263-8. PMID: 23187979.

59. Myers K, Vander Stoep A, Lobdell C. Feasibility of conducting a randomized controlled trial of telemental health with children diagnosed with attention-deficit/hyperactivity disorder in underserved communities. J Child Adolesc Psychopharmacol. 2013 Aug;23(6):372-8. PMID: 23952183.

60. Hogue A, Dauber S. Assessing fidelity to evidence-based practices in usual care: the example of family therapy for adolescent behavior problems. Eval Program Plann. 2013 Apr;37:21-30. PMID: 23314000.

61. Asarnow JR, Jaycox LH, Duan N, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. JAMA. 2005 Jan 19;293(3):311-9. PMID: 15657324.

62. Asarnow JR, Jaycox LH, Tang L, et al. Long-term benefits of short-term quality improvement interventions for depressed

youths in primary care. Am J Psychiatry. 2009 Sep;166(9):1002-10. PMID: 19651711.

63. Ngo VK, Asarnow JR, Lange J, et al. Outcomes for youths from racial-ethnic minority groups in a quality improvement intervention for depression treatment. Psychiatr Serv. 2009 Oct;60(10):1357-64. PMID: 19797376.

64. Wells KB, Tang L, Carlson GA, et al. Treatment of youth depression in primary care under usual practice conditions: observational findings from Youth Partners in Care. J Child Adolesc Psychopharmacol. 2012 Feb;22(1):80-90. PMID: 22251025.

65. Henggeler SW, McCart MR, Cunningham PB, et al. Enhancing the effectiveness of juvenile drug courts by integrating evidence-based practices. J Consult Clin Psychol. 2012 Apr;80(2):264-75. PMID: 22309470.

66. Liddle HA, Dakof GA, Henderson C, et al. Implementation outcomes of Multidimensional Family Therapy-Detention to Community: a reintegration program for drug-using juvenile detainees. Int J Offender Ther Comp Criminol. 2011 Jun;55(4):587-604. PMID: 20427547.

67. Weisz JR, Southam-Gerow MA, Gordis EB, et al. Cognitive-behavioral therapy versus usual clinical care for youth depression: an initial test of transportability to community clinics and clinicians. J Consult Clin Psychol. 2009 Jun;77(3):383-96. PMID: 19485581.

68. Bradshaw CP, Mitchell MM, Leaf PJ. Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. J Posit Behav Interv. 2010 Jul;12(3):133-48. PMID: 2010-13514-002.

69. Bradshaw CP, Pas ET, Goldweber A, et al. Integrating school-wide positive behavioral interventions and supports with tier 2 coaching to student support teams: The PBISplus model. Adv School Mental Health Promot. 2012;5(3):177-93. PMID: 2012-21616-004.

70. Kolko DJ, Baumann BL, Herschell AD, et al. Implementation of AF-CBT by community practitioners serving child welfare and mental health: a randomized trial. Child Maltreat. 2012 Feb;17(1):32-46. PMID: 22278087.

71. Gully KJ, Price BL, Johnson MK. Increasing abused children's access to evidence-based treatment: diffusion via parents as consumers. Child Maltreat. 2008 Aug;13(3):280-8. PMID: 18359928.

72. Garner BR, Godley SH, Dennis ML, et al. Using pay for performance to improve treatment implementation for adolescent substance use disorders: results from a cluster randomized trial. Arch Pediatr Adolesc Med. 2012 Oct;166(10):938-44. PMID: 22893231.

73. Glisson C, Hemmelgarn A, Green P, et al. Randomized trial of the Availability, Responsiveness, and Continuity (ARC) organizational intervention with community-based mental health programs and clinicians serving youth. J Am Acad Child Adolesc Psychiatry. 2012 Aug;51(8):780-7. PMID: 22840549.

74. Epstein JN, Langberg JM, Lichtenstein PK, et al. Use of an Internet portal to improve community-based pediatric ADHD care: a cluster randomized trial. Pediatrics. 2011 Nov;128(5):e1201-8. PMID: 22007005.

75. Epstein JN, Rabiner D, Johnson DE, et al. Improving attention-deficit/hyperactivity disorder treatment outcomes through use of a collaborative consultation treatment service by community-based pediatricians: a cluster randomized trial. Arch Pediatr Adolesc Med. 2007 Sep;161(9):835-40. PMID: 17768282.

76. Beidas RS, Edmunds JM, Marcus SC, et al. Training and consultation to promote implementation of an empirically supported treatment: a randomized trial. Psychiatr Serv. 2012;63(7):660-5. PMID: 22549401.

77. Lochman JE, Boxmeyer C, Powell N, et al. Dissemination of the Coping Power program: importance of intensity of counselor training. J Consult Clin Psychol. 2009 Jun;77(3):397-409. PMID: 19485582.

78. Carroll AE, Bauer NS, Dugan TM, et al. Use of a computerized decision aid for ADHD diagnosis: a randomized controlled trial. Pediatrics. 2013 Sep;132(3):e623-9. PMID: 23958768.

79. Lester H, Birchwood M, Freemantle N, et al. REDIRECT: cluster randomised controlled trial of GP training in first-episode psychosis. Br J Gen Pract. 2009 Jun;59(563):e183-90. PMID: 19520016.

80. Henggeler SW, Sheidow AJ, Cunningham PB, et al. Promoting the implementation of an evidence-based intervention for adolescent marijuana abuse in community settings: testing the use of intensive quality assurance. J Clin Child Adolesc Psychol. 2008 Jul;37(3):682-9. PMID: 18645758.

81. Wildman BG, Langkamp DL. Impact of location and availability of behavioral health services for children. J Clin Psychol Med Settings. 2012 Dec;19(4):393-400. PMID: 23053830.

82. Ronsley R, Rayter M, Smith D, et al. Metabolic monitoring training program implementation in the community setting was associated with improved monitoring in second-generation antipsychotic-treated children. Can J Psychiatry. 2012 May;57(5):292-9. PMID: 22546061.

83. Chadwick Center for Children and Families. Closing the quality chasm in child abuse treatment: Identifying and disseminating best practices. San Diego, CA: Author2004.

84. American Academy of Pediatrics, Committee on Quality Improvement, Subcommittee on Attention-Deficit/Hyperactivity Disorder. Clinical practice guideline: diagnosis and evaluation of the child with attention deficit/hyperactivity disorder. Pediatrics. 2000;105(5):1158 -70. PMID: 10836893.

85. American Academy of Pediatrics, Subcommittee on Attention-Deficit/Hyperactivity Disorder, Committee on Quality Improvement. Clinical practice guideline: treatment of the school-aged child with attention deficit/hyperactivity disorder. Pediatrics. 2001;108(4):1033-44. PMID: 11581465.

86. Henggeler SW, Halliday-Boykins CA, Cunningham PB, et al. Juvenile drug court: enhancing outcomes by integrating evidence-based treatments. J Consult Clin Psychol. 2006 Feb;74(1):42-54. PMID: 16551142.

87. Dennis M, Godley SH, Diamond G, et al. The Cannabis Youth Treatment (CYT) study: main findings from two randomized trials (Structured abstract). J Subst Abuse Treat. 2004;27(3):197-213. PMID: 15501373.

88. Chorpita BF, Regan J. Dissemination of effective mental health treatment procedures: maximizing the return on a significant investment. Behav Res Ther. 2009 Nov;47(11):990-3. PMID: 19632669.

89. Lester H, Allan T, Wilson S, et al. A cluster randomised controlled trial of patient-held medical records for people with schizophrenia receiving shared care. Br J Gen Pract. 2003 Mar;53(488):197-203. PMID: 14694695.

90. Okelo SO, Butz AM, Sharma R, et al. Interventions to modify health care provider adherence to asthma guidelines: a systematic review. Pediatrics. 2013 Sep;132(3):517-34. PMID: 23979092.

91. Gaynes B, Brown C, Lux LJ, et al. Relationship Between Use of Quality Measures and Improved Outcomes in Serious Mental Illness. Technical Brief No. 18. (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2012-00008-I.) AHRQ Publication No. 15-EHC003-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2015. www.effectivehealthcare.ahrq.gov/reports/final.cfm

92. James AC, James G, Cowdrey FA, et al. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev. 2015;2:CD004690. PMID: 25692403.

93. Wang PS, Lane M, Olfson M, et al. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen

Psychiatry. 2005 Jun;62(6):629-40. PMID: 15939840.

94. Collins KA, Westra HA, Dozois DJ, et al. Gaps in accessing treatment for anxiety and depression: challenges for the delivery of care. Clin Psychol Rev. 2004 Sep;24(5):583-616. PMID: 15325746.

95. Bryant J, Passey ME, Hall AE, et al. A systematic review of the quality of reporting in published smoking cessation trials for pregnant women: an explanation for the evidence-practice gap? Implement Sci. 2014;9:94. PMID: 25138616.

96. Rycroft-Malone J, Burton CR. Is it time for standards for reporting on research about implementation? Worldviews Evid Based Nurs. 2011;8:189-90.

97. Newhouse R, Bobay K, Dykes PC, et al. Methodology issues in implementation science. Med Care Res Rev. 2013;51(Suppl 2):S32-40.

98. Pinnock H, Epiphaniou E, Sheikh A, et al. Developing standards for reporting implementation studies of complex interventions (StaRI): a systematic review and e-Delphi. Implement Sci. 2015;10(1):42. PMID: 25888928.